

U-shaped behavior in multilingual children's written production of jokes

Eduard Abelenda Puigvert

Supervisor: Dr. Marta Fernández-Villanueva

Master's thesis

Applied linguistics and language acquisition in multilingual contexts

Universitat de Barcelona

2 February 2012

INDEX

Abstract	3
1. Introduction	4
2. Later language development and figurative uses of language	4
3. Joke telling in developmental perspective: comprehension and production	8
4. Joke telling and multilingualism	10
5. Hypothesis and research questions	11
6. Method	12
6.1 Corpus	12
6.2 Sample of participants	13
6.3 Procedure	13
7. Results	15
7.1 Speech acts	15
7.2 Text type	16
7.3 Structure	17
7.4 Expandability	18
7.5 Degree	20
7.6 Multimodality	21
7.7 Comments	21
7.8 Fragmentarity	22
7.9 Completeness	22
7.10 Punchline	22
7.11 Repetition	23
7.12 Combined errors	23
7.13 Display of Multilingualism	24
8. Discussion	25
9. Conclusion	30
References	31
APPENDIX	34

Abstract

The joke is a genre of oral literature acquired during childhood and adolescence. Together with other figurative uses of language it is bound to cognitive developments. The goal of this study is analyzing some developmental aspects of the written production of jokes and finding suitable coding categories for the analysis. A sample of 234 multilingual children and adolescents from 6 to 16 years of age were selected from the CesCa corpus (Catalan Written in Schools of Catalonia). They had been asked to write a joke or a funny situation. The categories proved quite suitable to code types of responses in terms of speech acts, types of texts, types of jokes, production errors and multilingualism in use. The results showed an increasing development by age groups and some u-shaped behaviors by age. We interpret u-shape behavior as a point of cognitive restructuring according to the Representational Redescription Model (RRM). Regarding multilingualism, jokes showed more code-switching than other humorous texts, but an increase of Catalan use and a descent in code-switching. The study of joke production in developmental perspective has proven to be a promising domain of inquiry deserving further exploration.

U-shaped behavior in multilingual children's written production of jokes

1. Introduction

During their early years, children are engaged in acquiring the main communicative functions and core grammar of their languages, i.e. they are engaged in becoming native speakers. However, there is still a long developmental path ahead of them until they become expert users of their language (Berman, 2004). Humor is one of the aspects believed to demonstrate a proficient command of a language and develops all through childhood and adolescence.

The goal of this study is finding some useful categories for analyzing different developmental aspects of the written production of jokes and ascertain whether developmental trends can be observed and explained. For that purpose, data was used from the CesCa corpus (Catalan Written in Schools of Catalonia): 234 texts produced by multilingual children aged 6 to 16. We first describe the theoretical framework of later language development, the Representational Redescription Model of Karmiloff-Smith (1992) and different developmental perspectives on child humor. Then we present the CesCa corpus and the categories that we used for data coding. Finally, we will present and discuss some results regarding the developmental trends we found.

2. Later language development and figurative uses of language

Later language development is the area of language acquisition that refers to the linguistic developments that occur beyond age five, which is considered as a sort of psycholinguistic frontier separating between early and later language development (Karmiloff-Smith, 1986 as cited in Tolchinsky, 2004).

While early language development centers on the acquisition of sentence structure, later language development goes beyond the sentence to acquire more control over discourse as an integrated whole. Already existing structures become more complex both in form and meaning. Not only structures change, but also functions develop as previously acquired forms are used for new functions and, conversely, old functions are expressed by new forms (Tolchinsky, 2004).

This period is also the time for the development of figurative speech and non-conventional meaning: metaphors, idioms, irony, ambiguity, lies, etc. The distinction between what is said and what is meant is refined by children and adolescents. The motor for this distinction is a cognitive process strongly linked to the capacity of conceiving different perspectives on reality. The

acquisition of the so-called theory of mind (ToM) underlies all these mental abilities. Having a theory of mind is having the capacity to take others' beliefs into account (Premack and Woodruff, 1978 as cited in Schlinger, 2009). Dialogue is both a consequence and a prerequisite of ToM. Literacy is also a sweeping feature of later language development. The developmental period coincides with school age, the time when all the linguistic repertoire related to written language and the literate society is acquired. Literacy increases the heterogeneity and complexity of language uses (Tolchinsky, 2004).

Tolchinsky (2004), highlights three important ages in this developmental period. First, age 5, around which children start using verbs expressing mental states with complement clauses and also start to comprehend the indirect nature of other's beliefs so that they can use language to deceive. However, Hale and Tager-Flusberg (2003, as cited in Schlinger, 2009) situate the emergence of ToM at age 4. Second, age 7, around which children distinguish and mark old and new information and start to process non-literal language. Third, age 11, at which children fully take into account other's internal states and intentions, understand some difficult idioms and cease relying on extralinguistic cues to interpret difficult surface or deep ambiguities. Tolchinsky (2004, p. 239) also states that only adolescents seem to produce figurative language creatively.

Figurative speech, theory of mind, dialogue and literacy are central to the interpretation of Jokes, since they do not reflect reality as it is, nor are they bona-fide communication, just like lies aren't. Telling a joke is proposing an alternative world. The joke teller must take into account the interpretations the hearer will make in order to choose the most suitable word combinations and fulfill the goal of the joke. Other aspects of later language development are present in joking. Language subtleties are exploited to create ambiguities in jokes, and discourse-internal relations and information management are crucial for success in the perlocutionary act of telling a joke.

The joke belongs to the domain of figurative language, since this domain includes uses of language for which what is said is not the same as what is meant. However, in jokes this is not in the same sense as in contextualized speech acts in which the propositional content is different from the illocutionary force and the perlocutionary act. Jokes are rather decontextualized and usually what is said leads to two compatible interpretations of the text that must be discovered, one more marked than the other (Attardo et al., 1994).

A joke is an anonymous genre of oral literature that can adopt the written mode as well. Jokes are usually called canned jokes because they are partially or completely recycled texts.

Prototypical jokes are narrative jokes that consist of a set-up in which a fictitious situation is provided. Since jokes are an economic genre, when circumstances do not play a role, the set-up is suppressed. Then, the real kernel of the joke is the dialogue. The narrative leads to an end where the text is closed by the punchline, which lends meaning to all preceding information. Anecdotes have a narrative resolution and may contain punchlines, but they do not close the text. When the receiver encounters the punchline, s/he is surprised and puzzled, stops the linear interpretation of the text and backtracks to check for other possible, more marked interpretations (Oring, 1989; Attardo and Chabanne, 1992; Attardo et al. 1994).

A mere formal description of the joke is not enough to distinguish it from other text types. A pragmatic point of view such as degree of context dependence can help in differentiating between joke and joking (Norrick, 2001 cited by Brock, 2008). Situational jokes must be suited to the communicative situation. They are improvised and appreciated precisely because they are not stereotypes. Anecdotes are proposed to the receiver as being authentic, with an emphasis on deeds and their consequences, while jokes emphasize linguistic and cognitive devices (Oring, 1989; Attardo and Chabanne, 1992; Attardo et al. 1994). An anecdote recounts something that happened to a character but a joke is more about what is happening to the receiver during the communication process. This could be the reason why the joke tends to stress the present tense and the anecdote tends to stress the past tense (Oring, 1989). Anecdotes and jokes will be differentiated in our data with the coding categories we propose. Moreover, types of jokes will also be differentiated (structures, types of second degree jokes etc.), together with errors in the construction of jokes (errors in set-up or punchline).

There are different competences that develop during childhood and adolescence or even during adulthood in a non-linear way. Behavior shows that a particular ability is performed for the first time, but after some time it is not performed, performed less or performed worse until it is later again performed with the same competence as the first time (Gershkoff-Stowe and Thelen, 2004). This has been observed for such different abilities as particular aspects of language development, artistic creativity, face perception, social cognition, phoneme discrimination, auditory localization or development of written representations. Since the seventies this phenomenon has been researched and called u-shaped behavior or u-shaped development (cf. Siegler, 2004).

Karmiloff-Smith (1992) develops a general theoretical model, called Representational Redescription Model (RRM), that accounts for cognitive development in different domains. The model also explains several u-shaped behaviors caused by cognitive development: language features (such as definite-indefinite article in French), face recognition, notation etc. She defends that apparent regression in behavior is in fact a constant progression in cognitive development. A u-shape indicates a moment in which an internal cognitive reorganization takes place.

The first behavioral mastery can occur because mental representations have been stored to produce a certain behavior effectively. Nevertheless, representations have been stored separately from other similar or related representations. This allows quick appropriate behavioral response. When stability is achieved, a new phase starts in which links between representations are done, redescribed and rewritten as new representations. The behavioral regression occurs in this phase because new representations take the form of reduced descriptions that lose details of the previous procedural information, and also because the weight of internal change prevents attention to external input. After that, a third phase starts in which internal representations created during the second phase are related again to external input.

Mandler (as cited in Karmiloff-Smith 1992, p. 21) proposes the following example: seeing a zebra as a striped animal and then associating it to a zebra crosswalk is something that only a species with the capacity of representational redescription would be able to do. The example could be seen as a suggestion of a relation between RRM, Metaphors and Jokes. Though, to our knowledge, no empirical data has been provided to link RRM to Jokes. Inasmuch as redescription brings together representations that were previously separately stored, it is possible to become aware of similarities that were not perceived before. According to Nerlich et al. (1998, p.360), "metaphor is a bridge between two incongruous conceptual domains" and operates by conceptual blending, while jokes operate by switching between conceptual domains to solve an incongruity. Certain jokes and metaphors are not understood until certain ages. The humorous devices that create and solve the incongruity of jokes might be related to a representational redescription.

3. Joke telling in developmental perspective: comprehension and production

Paul McGhee proposed four stages of humor development in childhood based on Piaget's model of child development (McGhee, 1979, cited by Martin, 2007). For each cognitive achievement described by Piaget, McGhee observes new possibilities for humor. At third stage, which covers ages 3 to 7, children start to cope with conceptual incongruity such as drawing a cat with two heads. The fourth and last stage is called "multiple meanings" and is reached around age 7, when children move from *preoperational stage* to *concrete operations stage* in Piaget's model. As Piaget believed, at this age, they become able to imagine the effect of an action without carrying it out physically. They also start to understand some forms of abstract humor that play on logical inconsistencies and require inferential thinking (Martin, 2007).

However, not all researchers agree about stages of humor development. Martin (2007) speculates about a possible fifth stage that could correspond to Piaget's formal operations stage, which begins at early adolescence. At this stage, thinking would be more abstract and more governed by logical reasoning than by experience. They are able to hypothesize logical sequences of actions and anticipate their consequences. Existential jokes about the meaning of life as well as second-degree jokes would be of interest to adolescents. Irony also particularly develops during adolescence, ironic compliments being appreciated after ironic criticism (Harris and Pexman, 2003, cited by Martin, 2007). Führ (2001), however, through interviews and surveys, discovers that humor in which adolescents are interested may differ from countryside to towns.

Nerlich et al. (1998) conducted a case study in which they observed that their subject produced simple and rehearsed riddles at the age of 5 years and one month, comprehended simple riddles at the age of 5 years and 5 months and fluently retold simple jokes and riddles at the age of 6. Nerlich et al (1998) is one of the few works that particularly considers production and the relationship between production and comprehension of humor. McGhee (1977, quoted by Nerlich et al. 1998, p. 356) states that an event is found funny only when the child feels confident about its impossibility or inappropriateness. Therefore, confidence is easier to feel when the incongruity is produced by oneself. This explains that self-generated incongruities appear before comprehension of incongruities. However, Nerlich et al (1998) observed that arbitrary incongruities keep being produced after meaningful incongruities are understood. That is one of

the reasons why comprehension studies have been given priority. However, the insight that production gives into cognitive processes cannot be attained equally with comprehension.

Ashkenazi and Ravid (1998) studied metalinguistic appreciation of linguistic humor and found that metalinguistic appreciation especially emerges from ages 7 to 9. At ages 10 to 11 there is a turning point with a much higher use of linguistic explanations, though still awkward. Clearly precise explanations with use of linguistic terms do not appear until high school, after age 13. Yalisove (1978) establishes a relationship between types of riddles and children's preference for them. Children were asked to write down their favorite riddle. Reality-based riddles are preferred by children aged 6 to 7, language ambiguity riddles are preferred by children aged 9 to 10, and finally, absurd riddles are preferred by adolescents aged 12 to 14.

The studies on joke acquisition focus more on comprehension than on production. Horgan (1981) already observed the lack of studies on joke production and an abundance of works about school age children's responses to humor with an emphasis on the role of cognition. However, unlike us, she focused on early age spontaneous jokes. Seventeen years later, Nerlich et al. (1998) pointed out that development of humor in production had yet to be well studied. Masten (1986) is interested in the relationship between cognitive competence and humor measured in different ways, one of which is production. However production is analyzed through gap filling in cartoons. Sanders (1986) points out that most previous child humor research on comprehension used a methodology consisting of two parts –production and comprehension–, because use of child-generated humor was preferred to adult-generated humor.

Corpus methods seem highly appropriate for the study of child joke telling. However, since production studies have been so scarce, corpora containing jokes told by children are scarce too. Lefort (1992) uses a corpus created by Cabaj et al. (1984, as cited in Lefort, 1992) containing jokes told by primary school children. At an initial stage of Lefort's (1992) experiment they analyze a portion of this corpus of jokes but collect their own data related to comprehension. However, they do not give details about criteria used and they do not analyze in detail either.

Baradel and Barbieri (2001) is a study that focuses on jokes told by children from grades 1, 3 and 5 (ages 6, 8 and 10, approximately). To our knowledge, this is the only study that analyzes competence in joke production. The researchers videotaped joke telling sessions and researched the distribution of pragmatic errors and content errors across ages. Presentation errors are those general errors that compromise the resolution of the incongruity by explaining it, or repeating an

expression too much, or temporarily interrupting the course of discourse. Content errors include forgotten passages or failed punchlines.

They found a u-shaped behavior in three categories of pragmatic errors: inopportune smiles, presentation errors and content errors. This study highlights the cognitive developmental importance of age 8 regarding joke telling, which is not usual in developmental studies of humor. Importance is also lent to age 10, since it is the age at which developmental gains appear. Something important lacking in Baradel and Barbieri (2001) is an interpretation of the findings using some sort of previous theoretical explanations of the u-shape. They provide a hypothetical explanation for their findings with no reference to previous research.

4. Joke telling and multilingualism

Bilingualism is usually exploited for humor in bilingual societies throughout the world. It is a tool to create bilingual Jokes and puns but also to develop a personal or collective identity. Since the uses of humor are different throughout societies it is considered that when learning a language, one of the last language aspects acquired are those relating to the comprehension and use of humor (Vaid, 2006). Simarro (2011, December) found that even level C L2 learners of Spanish could not understand polysemy-based jokes.

Although Catalan is always in a language contact situation in all the territories where it is spoken, there has been no research regarding humor in Catalan as part of a multilingual linguistic ecosystem. Actually humor and joke telling in Catalan has scarcely been researched even in monolingual perspective. As far as we know, Arts (1988) conducted a doctoral dissertation on jokes, their linguistic and rhetoric devices and their relation to literature. Oriol (2000) published a study with a folklore perspective about joke telling and identities in Andorra. Viana (2004) addresses humor as a social and discursive phenomenon with some references to jokes. Reyes et al. (2009) published an article based on the CesCa corpus where they applied different techniques of automatic humor recognition to differentiate non-humorous narratives from jokes. It is the only study published based on jokes from the CesCa corpus.

Bilingualism and code-switching is inherent in contemporary joke telling practice in Catalonia. Telling a joke in Spanish is not seen as an identity switch (Pujolar, 1997, p. 27), because code-switching in jokes is interpreted differently from joking. The subjects in our study speak languages other than Catalan or Spanish at home, and many of them arrived in Catalonia as

immigrants. Thus, they had to learn both Catalan and Spanish at the same time, together with the cultural implications of humor use and the way in which code-switching is used in jokes. Learning new ways of using humor in a new language might change even what one finds funny (Vaid, 2006).

We know of only one developmental study of humor that takes into account bilingualism and home language: Sanders (1986), a study about humor production and comprehension in American Sign Language (ASL). They found a remarkable difference between children with deaf parents (native speakers of ASL) and children with hearing parents (non-native speakers of ASL).

The fact that the children and adolescents that we selected from the CesCa corpus for this study learned Catalan later does not necessarily mean that they master Catalan poorly. Tolchinsky et al. (submitted) investigated lexical development operationalized as word length, use of adjectives and nominalizations. They found that participants who speak other languages at home and have known Catalan for more than four years score slightly better than participants who only speak Spanish at home. We cannot reject that something similar could happen regarding joke telling or other figurative uses of language.

5. Hypothesis and research questions

We hypothesized that there might be u-shaped behavior in children's performance of written jokes. We based our hypothesis on findings made by Baradel and Barbieri (2001), which reveal a u-shaped distribution in the production of jokes told by first-grade, third-grade and fifth-grade schoolchildren. When we observed our data superficially, some u-shapes seemed to appear regarding frequency of joke telling.

1. What is the distribution of text types across ages? Do we find any u-shaped distribution in jokes?
2. Which type of joke increases or decreases with age? Is the type of joke (structure, expandability, degree, multimodality) parallel to the distribution of text-type jokes?
3. When do errors primarily occur? Do errors primarily occur at the age when joke production decreases or at the previous age?
4. What is the distribution of languages in the texts?
 - a. Are jokes different from other texts regarding language use?

- b. Does Catalan increase with age while Spanish and Code-switching decrease?
- c. What is the proportion of free code-switching and code-switching constrained by the joke?
- d. How are translation errors distributed?

6. Method

6.1 Corpus

Data from the corpus CesCa was analyzed to answer to the research questions. It is publicly available at <http://clic.ub.edu/cesca>. For a extended description of the corpus see Llauradó et al. (submitted).

The CesCa Corpus is a reference corpus of Catalan written at school in Catalonia. These writing productions were gathered in 2006 with the participation of 2,161 children and teenagers, who carried out the tasks during their Catalan lessons at school. Through a sociolinguistic questionnaire, information was gathered on the participants. The home language of 16 % of the participants was a language other than Catalan or Spanish.

The corpus contains vocabularies (lists of words) in different semantic fields. It also contains texts of different genres: narrative texts, argumentative texts, definitions and jokes. All texts have been morphologically analyzed and lemmatized. The 11,332 texts yield 207,028 tokens, 131,263 types and 113,160 different lemmas.

There are 1867 texts under the label “jokes”. Jokes contain 58,561 tokens, 42,326 types and 38,229 lemmas. The mean length of a joke is 27.10 words. Jokes were added to the corpus as a representation of colloquial language use, closer to orality. For this reason the instruction was: “Explica un acudit que t’agradi o una situació que t’hagi fet riure molt” (Tell a joke you like or a situation that made you laugh a lot).

Although participants were asked to write in Catalan, 314 texts from the CesCa corpus were completely written in Spanish. Of these, 162 related to the explanation of a joke. Tolchinsky and Llauradó (submitted) suppose that the language in which the joke is usually told may have motivated the language choice. Jokes had more words from other languages than other types of text. They explain this arguing that jokes are highly associated with orality and therefore not much restricted in terms of code-switching.

Despite the fact that classrooms are highly multilingual environments, non-Catalan forms accounted for 4% of the total number of lexical forms in the corpus. The context of elicitation was considered as the main explanatory factor for the scarcity of non-Catalan forms in the discourse production of the children (Llauradó, submitted).

6.2 *Sample of participants*

We selected 234 informants for this study, all of them speaking a language other than Catalan or Spanish at home. These subjects are, in general, more multilingual than their peers, since they are exposed to Catalan, Spanish and English at school, like every pupil, but they also speak another language at home. We selected them because we are interested in the way they display multilingualism.

Their languages might belong to very different families. During the collection of the corpus it was asked how long each subject had been in contact with Catalan. In our research, we did not analyze the influence of each language or the time of exposure to Catalan, but we did take into account the way multilingualism and code-switching was displayed in jokes.

Among 291 subjects that speak other languages at home, we selected 234 for our study, because we were interested in three age groups: 6 to 8-year-olds, 10 to 12-year-olds and 14 to 16-year-olds. A relative decrease in the production of jokes in comparison to other texts labeled as jokes for 7-year-olds, 11-year-olds and 15-year-olds drew our attention because it indicated a possible u-shape. Thus, 9-year-olds and 13-year-olds were left out.

6.3 *Procedure*

We used a qualitative approach to make coding categories emerge from data with the guide of previous research (Oring, 1989; Hetzron, 1991; Attardo and Chabanne, 1992; Attardo et al., 1994; Ritchie, 2004). We identified different text types under the label “jokes”. Hence, in the category “speech act” we decided to first classify the texts according to the pragmatic interpretation subjects made of the instructions they were given. According to their respective interpretations, they produced different types of text. In an appendix, we extended the presentation of categories and codes we present now, together with examples and coding criteria.

The category *text type* helps answer the first research question. It specifies the previous category and distinguishes between two types of directive interpretations: jokes and

anecdotes/tales; and two types of assertive interpretations: more narrative assertive (quasi-anecdotes) and less narrative assertive (more descriptive).

The categories *structure*, *expandability*, *degree*, *multimodality* and *comments* were created to classify jokes and thus help answer the second research question. The category *structure* refers to the superficial structure of jokes: narrative jokes, riddles or shortest narratives. To refine the broad codes we used for structure, we created the category *expandability*. Jokes that can be varied and expanded without compromising the humor are more expandable, rewordable jokes are less expandable and fixed jokes are not expandable. The category *degree* differentiates second-degree from first-degree jokes, and it also differentiates between several types of second-degree jokes. *Multimodality* serves to classify jokes according to requirement of non-verbal or para-verbal support. The *comments* category accounts for added text besides the joke.

To answer to the third research question we used certain categories of errors: fragmentarity, completeness, punchline, repetition, and the code *translation failure*, which belongs in the category *display of multilingualism*. To describe an error in the joke we compared participants' productions to a model found on Internet of the same joke told in a successful way. Almost all the jokes in the corpus were identifiable. *Fragmentarity* refers to unended texts. *Completeness* refers to elements lacking in the set-up of the joke. *Punchline* refers to absent or failed punchlines. *Repetition* only codes errors in repetition-based jokes. *Translation failure* occurs when the joke is not understood because the joke teller is not aware that it requires switching into Spanish and translates everything into Catalan. All categories related to errors will be used to create an index of errors that allows us adding all the errors and compare them across ages.

Finally, the fourth research question will be answered by means of the category *display of multilingualism*. It includes a code for texts in Catalan, Spanish and different types of code-switching: mainly free code-switching and code-switching required by the joke, i.e. bilingual jokes or different languages in punchline and set-up.

7. Results

In the results section we include the distribution of codes from coding categories crossed by age group or age. Age groups show us a developmental overview, while age shows us the internal diversity of age groups and can reveal u-shapes.

Research Question 1: What is the distribution of text types across ages? Do we find any u-shaped distribution in jokes?

To answer the first research question, we observed the frequency of text types, first classifying them according to the speech acts that were interpreted in the instruction and then according to text types such as jokes, anecdotes, more narrative assertive and less narrative assertive (descriptive). We indicated the percentages of each code vis-à-vis the total, which is the whole group of subjects considered at the moment (age group or age). The consideration of each age allowed possible u-shapes to be detected inside an age group.

7.1 Speech acts

Table 1. Distribution of speech acts by age (in percentages)

	Group 1			Group 2			Group 3		
	6 n=13	7 n=16	8 n=23	10 n=37	11 n=36	12 n=32	14 n=31	15 n=23	16 n=23
Jokes	61,5	31,3	47,8	54,1	47,2	65,6	64,5	34,8	73,9
Anecdotes	15,4	6,3	4,3	18,9	16,7	15,6	6,5	21,7	8,7
+Narrative	0	6,3	13	10,8	16,7	6,3	9,7	4,3	8,7
-Narrative	7,7	25	21,7	8,1	13,9	6,3	6,5	21,7	4,3
Other	15,4	31,3	13	8,1	5,6	6,3	12,9	17,4	4,3

Regarding the speech act that subjects performed in their response, our results per age group show that frequency of directive increases with age and eventually stabilizes (group 1: 59.6%; group 2: 75.2%; group 3: 74%). Assertive interpretations, on the other hand, decrease slightly (group 1: 28.8%; group 2: 22.9%; group 3: 19.5%). In table 1 we observe that directive is especially frequent in six-year-olds, while seven-year-olds and eight-year-olds produce much less of them. In group 2 the frequency of directives is quite high, but 11-year-olds produce much less

than the others, thus forming a u-shaped distribution. We observe that values for assertive mirror those of directive. Regarding the category “other”, frequency is especially high for group 1 in general. However, these percentages are based on a small number of elements.

7.2 Text type

The results for text types are the following: production of jokes increases (group 1: 46.2%; group 2: 55.2%; group 3: 58.4%). On the other hand, anecdotes are more frequent in group 2 (group 1: 7.7%; group 2: 17.1%; group 3: 11.7%). Likewise in group 2, narrative assertive (quasi-anecdotes) are more frequent (group 1: 7.7%; group 2: 11.4%; group 3: 7.8%). The curve has the same shape. Less narrative assertive are quite high in the first age group, but then decrease (group 1: 19.2%; group 2: 9.5%; group 3: 10.4%). Finally, the category “other” also has a high frequency for group 1 and a lower one for the rest, especially group 2 (group 1: 19.2%; group 2: 6.7%; group 3: 11.7%).

Table 2. Distribution of text types by age (in percentages)

	Group 1			Group 2			Group 3		
	6 n=13	7 n=16	8 n=23	10 n=37	11 n=36	12 n=32	14 n=31	15 n=23	16 n=23
Jokes	61,5	31,3	47,8	54,1	47,2	65,6	64,5	34,8	73,9
Anecdotes	15,4	6,3	4,3	18,9	16,7	15,6	6,5	21,7	8,7
+Narrative	0	6,3	13	10,8	16,7	6,3	9,7	4,3	8,7
-Narrative	7,7	25	21,7	8,1	13,9	6,3	6,5	21,7	4,3
Other	15,4	31,3	13	8,1	5,6	6,3	12,9	17,4	4,3

In table 2 we can observe that, by age, Jokes present the most salient distribution. The lowest percentages are to be found in ages 7, 11 and 15, which are the central ages in each of the three age groups. Therefore, there is a u-shaped distribution inside each age group. Anecdotes present a more irregular distribution. *More narrative assertive* increase inside group 1, stabilize along ages 8, 10 and 11, but are similarly low from age 12 onwards. *Less narrative assertive* appears frequently at ages 7 and 8, are low at ages 12 and 14, but surprisingly we find the highest value at age 15, just before finding the lowest at age 16.

Group 1 is moderately homogeneous. For 7 and 8-year-olds, jokes are compensated by assertive, but anecdotes are insignificant. Group 2 is more internally homogeneous than the others. Anecdotes and assertive play a role in group 2. Group 3 is not internally homogeneous. At age 14 and 16 anecdotes and assertive have a similarly low presence, while at age 15 they have a high presence. This emphasizes the difference of age 15.

When the production of jokes descends at a certain age, what text types increase? Increases in these moments are different for every code. At age 15 all categories change frequency remarkably vis-à-vis previous and following ages.

Research Question 2: Which type of joke increases or decreases with age? Is the type of joke (structure, expandability, degree, multimodality) parallel to the distribution of text-type jokes?

Through the description of the categories of structure, expandability, degree and multimodality, we attempt to ascertain the frequencies of joke types. We can then compare the results with the previous description, in which we saw a u-shaped distribution of joke quantity.

7.3 Structure

Narrative jokes comprise the majority of jokes told (60.6%; n=127). However, when we compare the three age groups, they decrease (group 1: 83.3%; group 2: 67.2%; group 3: 40%). The descent between group 2 and group 3 is the biggest. Riddle-like jokes are the smallest group of jokes (16.5%; n=127) and they increase through age groups (group 1: 4.2%; group 2: 15.5%; group 3: 24.4%). The frequency of shortest narrative jokes is somewhat higher (22.8%; n=127) and increases by age group (group 1: 12.5%; group 2: 17.2%; group 3: 35.6%). The distribution of joke structures across age groups displays a complementarity between narrative jokes and the other categories.

Table 3. Distribution of joke structures by age (in percentages)

	Group 1			Group 2			Group 3		
	6 n=8	7 n=5	8 n=11	10 n=20	11 n=17	12 n=21	14 n=20	15 n=8	16 n=17
Narrative	87,5	80	81,8	75	58,8	66,7	40	37,5	41,2
Riddles	0	0	9,1	10	11,8	23,8	20	12,5	35,3
Shortest	12,5	20	9,1	15	29,4	9,5	40	50	23,5

In table 3 we can see more precisely how these joke structures are distributed. For narrative jokes, group 1 and group 3 are quite homogeneous, in contrast with group 2. Age 11 offers the lowest percentage of this structure inside group 2, forming a slight u-shape. For Riddles, group 1 is almost inexistent. Riddles start being told at age 8 and remain stable through age 10 and 11, then increase at age 12. Group 3 is the least homogeneous. Shortest narrative jokes are told at all ages, although there are very few examples in group 1. They present a kind of parallel to frequency of jokes, because at every age group the highest frequency is the one in the middle (ages 7, 11 and 15).

7.4 Expandability

The expandability category, when observed by age groups, presents the same curve shape as the structure category. Thus, jokes that can be expanded decrease by age group, whereas jokes that can be reworded and fixed jokes increase. There is a clear relation between them, because narrative jokes can usually be expanded, riddles can only be reworded or fixed and shortest narratives cannot be changed at all.

Table 4. Distribution of expandability in jokes by age (in percentages)

	Group 1			Group 2			Group 3		
	6 n=8	7 n=5	8 n=11	10 n=20	11 n=17	12 n=21	14 n=20	15 n=8	16 n=17
Widening	50	60	54,5	55	29,4	52,4	25	25	29,4
Rewording	25	20	18,2	20	41,2	33,3	35	25	47,1
Fixed	25	20	27,3	25	29,4	14,3	40	50	23,5

In table 4 we can observe, first of all, that all categories have some presence at all ages. We also observe some u-shaped and inverted u-shape distributions. We find an increase in expandability at age 7 and a decrease at age 11.

When we compare the total number of jokes in each structure and each expandability category, we find different distributions despite the correspondences between the categories: structure (narrative: 60.6%; riddle: 16.5%; shortest: 22.8%; n=127), expandability (expandable: 40.9%; rewordable: 31.5%; fixed: 27.6%; n=127). This means that inside each structure there is, in fact, a diversity that the expandable category can partially account for.

When we subtract values in expandability from values in structure we obtain values that provide interesting information. Subtracting expandable from narrative structures yields values, which mean that there are narrative jokes that cannot be expanded. At least 19.7% are non-expandable narrative jokes, since 40.9% of total expandable jokes subtracted from 60.6% of narrative jokes yields said 19.7% as a result. If we proceed similarly with the pair “riddles” and “rewording” and the pair “fixed” and “shortest”, we obtain mainly negative values which mean that rewording comes not only from riddles but also from narratives. For the other pair it means that some fixed jokes come from riddles and narrative jokes. After the subtraction between “fixed” and “shortest”, what we find remarkable is that for ages 7, 11 and across group 3, the result of subtraction is 0, which means that all shortest jokes are fixed, all riddles are in the rewordable category and narratives are partly in narratives and partly in rewordable. It is important to note that non-expandable narrative jokes decrease by age group similarly to narrative jokes in general and the expandability category in general. After subtraction, the numbers for narrative jokes that are not expandable would be roughly: (group 1: 29%; group 2: 21%; group 3: 13%).

When we compare the results of structure and expansion by age, we find differences and similarities regarding the distribution of values. We can observe in table 5 that expandability has u-shapes in groups 1 and 2, and after subtraction we can see that non-expandable narrative jokes have u-shapes there too. On the other hand, narrative jokes do not display this same u-shape exactly. We find a coincidence in group 1 of an inverted u-shape for shortest and a u-shape for fixed. We find another coincidence in group 2: a u-shape for both narrative and expandability. The

most salient coincidences are found in group 3, which follows the same pattern all along both categories.

7.5 Degree

First-degree jokes are the great majority (92.1%; n=127). The descent of first-degree jokes at group 2 highlights the importance of second-degree jokes for our oldest subjects (group 1: 95.8% n=24; group 2: 87.9% n=58; group 3: 88.9% n=45). The percentages of first-degree jokes by age are as follows (age 6: 100%; age 7: 100%; age 8: 90.9%; age 10: 85%; age 11: 88.2%; age 12: 90.5%; age 14: 90%; age 15: 87.5%; age 16: 88.2%). We observe that second-degree jokes gain importance from ages 8 to 14.

Table 5. Distribution of second-degree jokes by age (absolute numbers)

	Group 1			Group 2			Group 3			Total
	6 n=8	7 n=5	8 n=11	10 n=20	11 n=17	12 n=21	14 n=20	15 n=8	16 n=17	
Particular	0	0	0	1	0	0	1	0	0	2
Type	0	0	0	1	0	0	0	0	0	1
Concept	0	0	1	1	2	2	1	0	0	7
Total	0	0	1	3	2	2	2	0	0	10

The type of second-degree joke written the most is the one based on the very concept of joke. It is also the first one to appear and the only one present at all ages that register second-degree jokes. It is followed by jokes that play on a particular joke and finally those that play on a type of joke. In our data, second-degree jokes seem especially important at age 10, because all types of second-degree jokes are discovered by then. However, since the number of examples is so small, we can hardly distinguish among the ages at which we find second-degree jokes.

7.6 Multimodality

Table 6. Distribution of multimodality forms by age group (in percentages)

	Group1	Group 2	Group 3
	n=24	n=58	n=45
Verbal	45,8	94,8	86,7
Non-verbal	8,3	1,7	0
Para-verbal	45,8	3,4	13,3

The percentage of jokes requiring gestures or a particular voice performance are primarily found in group 1. Gestural jokes decrease as age increases. Jokes requiring voice only decrease in group 2 but increase again slightly in group 3. The u-shape is thus very moderate. Neutral jokes increase especially at group 2 as a consequence of a decrease of the other categories. Gestural jokes are very scarce: one example at age 6, one at age 8 and one at age 11. There are slightly more voice jokes, though altogether also very few.

7.7 Comments

We also coded comments added to the main text. Comments increase at every age group (group 1: 1.9%; group 2: 3.8%; group 3: 20.8%). However, they are concentrated at group 3, where we find the following age distribution (age 14: 25.8% n=31; age 15: 17.4% n=23; age 16: 17.4% n=23). First comments occur at age 7, and especially increase at age 10.

Research Question 3: When do errors primarily occur? Do errors primarily occur at the age when joke production decreases or at the previous age?

To answer this question, we will consider the frequency of some errors coded by means of the categories of fragmentarity, completeness, punchline and repetition. Then we will combine all the errors to see whether the frequencies of distribution change. Finally we will compare the distribution with that of joke frequency by age.

7.8 Fragmentarity

The percentage of joke fragments decreases by age group (group 1: 20.8% n=24; group 2: 5.2% n=58; group 3: 2.2% n=45). The absolute number of joke fragments is very low and is distributed across ages as follows (age 6: 2; age 7: 2; age 8: 1; age 10: 1; age 11: 1; age 12: 1; age 14: 1; age 15: 0; age 16: 0). Fragments do not disappear until age 15.

7.9 Completeness

Three of the codes for the category “completeness” have the highest percentages: complete (78%), almost complete (7.9%) and incomplete (8.7%). The rest of the categories representing errors have very few occurrences. The percentages of the code “complete” increase by age group (group 1: 45.8% n=24; group 2: 82.8% n=58; group 3: 88.9% n=45) while the codes for errors decrease: almost complete (group 1: 12.5%; group 2: 8.6%; group 3: 4.4%) and incomplete (group 1: 20.8%; group 2: 6.9%; group 3: 4.4%). The only group that registers a presence of all kinds of errors gathered under “completeness” is group 1.

Inside each age group we also find a progression by age, with one exception in group 2 at age 12 (age 10: 80%; age 11: 88.2%; age 12: 81). This means that at age 12, errors regarding completeness increase. The rest of errors will be considered under the section called “combined errors”, in which we sum up all errors.

7.10 Punchline

When we consider age groups, the results are parallel to those in completeness, since correct punchlines increase by age group (group 1: 70.8%; group 2: 84.5%; group 3: 93.3%), while failed and missing punchlines decrease. Failed punchlines (group 1: 12.5%; group 2: 10.3%; group 3: 4.4%). Missing punchlines (group 1: 16.7%; group 2: 5.2%; group 3: 2.2%). Failed punchlines are in 8.7% of total jokes, while missing punchlines are in 6.3%. When we observe the correct results by age, we find the following distribution (age 6: 62.5%; age 7: 40%; age 8: 90.9%; age 10: 90%; age 11: 76.5%; age 12: 85.7%; age 14: 90%; age 15: 100%; age 16: 94.1%), i.e., a decrease at ages 7 and 11 and an increase at age 15. There are errors regarding punchline at all ages except for age 15.

7.11 Repetition

Jokes that require repetition are 12.5% (n=127) of the total, distributed at a rate of 3 per age at most. They are found at all ages except 14 and 16. The proportion of jokes with repetition decreases as age rises. This means that repetition jokes, which are narrative, are more typical of group 1 and 2 while adolescents from group 3 lend them less importance. The biggest descent is found between age 8 (27.3%) and age 10 (15%). Repetition errors (“same” and “no repetition” types) are found at ages 6, 8, 10, 11 and 12, but no error is found in group 3. All jokes with repetition at age 7 and age 15 display successful repetitions.

7.12 Combined errors

When the errors that we coded in different categories are combined in absolute values, there are different results. To assign the errors a significance relative to the total jokes told at every age, we created an index dividing the total errors by the total jokes at each age and then multiplying the result by 1000 to avoid decimals. Such an index is simple to create because there is only one code (or two in a few cases) per subject and text. We include translation errors that were coded in the multilingualism category discussed later.

Table 7. Distribution of combined errors expressed by an index

	Group1			Group 2			Group3		
	6 n=8	7 n=5	8 n=11	10 n=20	11 n=17	12 n=21	14 n=20	15 n=8	16 n=17
Fragment	2	2	1	1	1	1	1	0	0
Almost	0	1	2	2	1	2	1	0	1
Incomplete	2	1	2	2	1	1	1	1	0
Takeups	0	1	0	0	0	0	0	0	0
Character	0	0	1	0	0	0	0	0	0
Failed punchline	2	1	0	1	3	2	2	0	0
Same	0	0	0	1	0	1	0	0	0
No repetition	2	0	2	0	1	0	0	0	0
Translation	0	0	1	0	1	1	2	0	0
Total	8	6	9	6	8	8	7	1	1
Error Index	1000	1200	818	300	470	380	350	125	58

Our results here suggest an inverted u-shape distribution of errors with the highest value inside the groups at age 7 and 11. This coincides with a decrease of joke production as well. Group 3 is a different case, since there is a decrease in joke production at age 15 but a decrease of errors at every age.

The fourth research question deals with language choice. We will look at how language choice is distributed mainly in jokes, but also comparing them with other types of text. We expect that Spanish and code-switching decrease, but has yet to be confirmed by data.

Research Question 4: What is the distribution of languages in the texts?

- a. Does Catalan increase with age while Spanish and code-switching decrease?
- b. Are jokes different from other texts regarding language use?
- c. What is the proportion of free code-switching and joke-constrained code-switching?
- d. Translation failure (joke not understood because of translation failure).
- e. Are there any parallels between the display of multilingualism and errors or types of text?

7.13 Display of Multilingualism

Display of multilingualism is distributed differently depending on whether we consider only jokes or the rest of the texts. Only jokes (Catalan: 77.2%; Spanish: 11%; code-switching: 11.9% no text: 0%). The rest of the texts (Catalan: 84.1%; Spanish: 10.3%; code-switching: 2.8%; no text: 2.8%). In jokes, Catalan is less present but code-switching is more present, whereas Spanish is almost invariable. Code-switching in jokes is free in 7.9% of cases and constrained by the joke in 4% of cases (bilingual jokes and joke demands). In the other texts, code-switching is free in every case.

Table 8. Distribution of languages used in jokes (in percentages)

	Group 1			Group 2			Group 3		
	6	7	8	10	11	12	14	15	16
	n=8	n=5	n=11	n=20	n=17	n=21	n=20	n=8	n=17
Catalan	50	60	81,8	65	76,5	90,5	75	75	94,1
Spanish	12,5	20	9,1	10	11,8	9,5	20	0	5,9
Code-switching	37,5	20	9,1	25	11,8	0	5	25	0

Jokes told completely in Catalan increase by age group (group 1: 66.7%; group 2: 77.6%; group 3: 82.2%) and also inside each age group, as we can see in table 2. Each age group starts with a lower percentage than that of the end of the previous group. Jokes told in Spanish do not vary much across age groups (group 1: 12.5%; group 2: 10.3%; group 3: 11.1%). Table 3 shows how at age 7 and age 14, Spanish increases remarkably. Code-switching, on the other hand, decreases (group 1: 20.8%; group 2: 12%; group 3: 6.6%). Table 3 also shows how, inside each group, code-switching decreases as well, except in group 3 where, internally, there is a clear inverted u-shape. Parallel to the use of Catalan, code-switching starts in group 2 with a higher frequency than it ends in group 1. The same happens in non-joke texts.

Code-switching when the joke demands it has one occurrence in some ages, once in each age group. Bilingual jokes occur only once at age 6 and once at age 8. Translation failures also have few occurrences but, surprisingly, occur primarily at ages 8 to 14.

8. Discussion

We asked ourselves which was the distribution of text types across ages and whether a u-shape could be found. To ascertain the different text types that the instructions elicited and see whether a developmental trend could be observed, we categorized the text types children produced. A majority of the texts were directive responses, as was preferred by the corpus designers. As expected, directives increased with age and assertive decreased. Although Baradel and Barbieri (2001) asked their subjects for jokes only, like us, they found an increase of jokes with age.

However, in our data there is more clearly a decision subjects are making about what is being requested in the instructions: is it information or production of a text that causes hilarity? There seems to be a social convention about interpreting these instructions as seeking language productions that are funny and not seeking a description of what makes one laugh. As they grow older, subjects tend to the former.

The decrease of directives at age 7, 11 and 15 coincides with a decrease in the production of jokes, although by age group, directives increase. This reveals a u-shaped distribution of the quantity of directives and jokes produced, in relative terms.

The fact that less narrative assertive decrease while anecdotes and more narrative assertive peak in group 2 points to a developmental trend: less narrative assertive are gradually

abandoned. However, we observe that less narrative assertive increase precisely when jokes decrease, mirroring the u-shape. We still do not know why less narrative assertive might influence jokes, but there is a possible relationship. Age 15 is different from the rest because it is the peak of strong or weak u-shapes insofar as distribution of all text types, but we cannot ascertain why this occurs with the information we have now.

According to Gershkoff-Stowe and Thelen (2004), u-shaped behavior is a decrease in performance that can be quantitative, qualitative or both. Thus far, we have found at least a u-shaped distribution of joke quantity. Baradel and Barbieri (2001) found a qualitative difference but not a quantitative one.

We wondered about the types of joke that could increase with age and also about possible parallelisms between quantity of jokes and types of jokes. The analysis through coding categories clarified that narrative jokes are the majority of jokes and they do not exactly parallel the u-shaped distribution of quantity of jokes by age. However, we can say that at age 7 and 11 narrative jokes decrease considerably. Shortest jokes, however, do parallel the u-shape of joke frequency. Riddles also register a u-shape at age 15.

On the other hand, when we observe these types of jokes across age groups we find that narrative jokes decrease by age group while riddles and shortest jokes increase. Apparently, narrative jokes occupy the bulk of joke production of youngest children, but as they grow, they gradually introduce the other types. Nevertheless, the significance of shortest jokes at older ages is surprising, because usually they are based on a very simple semantic resource. Perhaps the same reluctance that leads them not to tell jokes at certain ages causes them to choose these easier jokes.

The introduction of other types of jokes, even those that we do not distinguish in the present work, might be guided by the relationship between cognitive development and the linguistic and humorous devices in the joke (Martin, 2007, p.238). Nevertheless, we do not see this cognitive demand in the shortest jokes of our sample (such as “*va un peix i s’ofega*” - there was a fish and it drowned). We must add, however, that not all shortest jokes are so simple, because some of them become so stereotyped and repeated that adolescents like to play with their structure. The original is “*van dos i el del mig cau*” (two [people] were walking and the one in the middle fell down) and a retelling by a 16-year-old is: “*Van dos en una moto i el del 1/2 cau per la*

finestra del darrera” (Two [people] were riding on a motorbike and the one in the middle fell out the back window). Then the joke becomes a second-degree joke, that is, an advanced appreciation of creativity (cfr. Attardo and Chabanne, 1992, p. 174).

In our results, second-degree jokes do not become significant until age 10, when we find examples of every type of second-degree joke we coded. These jokes appear as a characteristic of group 2 and group 3 but not group 1. The categories Multimodality and Comments also help us describe group 2 and 3 together, since both point to a development that starts later. On the other hand, the majority of jokes that would require gestural or voice support are found at youngest ages. Children at these ages try to write the jokes they choose although they need gestural or auditory support to be effective with them. It is an inappropriateness that highlights the direct link between literacy skills and the ability to write a joke, which is a genre of oral literature.

We also see the possibility of a cognitive demand in some riddles, such as this one told by a 16-year-old: “En què s'asemblen un tomàquet i una patata? Les dues són vermelles menos la patata” (Why is a tomato like a potato? Both are red except for the potato). Riddles start being told by our 8-year-old subjects and thereafter frequency increases. Therefore, in our data, riddles are a feature of older subjects. In their case study, however, Nerlich et al. (1998) found that riddles were told at the age of 5 years 5 months. Also, Yalisove (1978) stated that riddles based on conceptual tricks related to reality are more popular among 6 or 7 year-olds and riddles based on absurdity are preferred by teenagers. So why is it that we found no riddles at ages 6 or 7? It must be taken into account that some jokes can be transformed into narrative form or into riddle form and there might be a cultural factor in the choice: perhaps Anglo-Saxon cultures cultivate the riddle more.

Expandable jokes present a u-shape, as do wholly non-expandable narrative jokes. The category Expandable jokes actually includes several different superficial structures and humorous devices, while the concept of non-expandable narrative jokes points to a non-prototypical form of narration, which is the most reproductive and least creative. Furthermore, it selects certain humorous devices or textual structures such those of scalar humor (Bergen and Binsted, 2004). Therefore, the u-shape seems to be linked more with certain humorous devices (that may favor expandability or not) than with narrativity. In future studies, we should segment narrative jokes and riddles into different superficial structures and humorous devices which are narrower in order to reveal their age prevalence and order of acquisition. In any case, the u-shapes we observed out

of the deductions from the interplay of expandability and “narrative” mean that some humorous devices are being abandoned while others are being adopted. Thus, the reduced joke production is related to a change in the type of jokes that children are acquiring at ages 7 and 11. The general coincidences regarding distributions of structure types and expandability types in group 3 might be related to the development of some humorous devices but we cannot prove it. Narrower categories might also help discover what actually happens after age 12.

We wondered about the possible coincidence of a u-shape in quantity of jokes and another u-shape regarding errors. When we observe the results by age groups, we find that all categories general to jokes reveal a linear and steady development: fragments decrease, incomplete and near-complete jokes decrease, and failed or missing punchlines also decrease. These results indicate the subjects’ general control over the jokes they choose to write.

We decided to combine all errors by adding them. We created an index assigning to errors a value relative to the total number of jokes produced at a particular age. We observed that at a particular age, while a category of errors could indicate quite a lot of errors, another category could indicate none. Therefore, the sum of errors is pertinent. The final result of combined errors indicates an increase of errors at ages 7 and 11. This implies a u-shape regarding performance at a quality level that coincides with a u-shape at the quantity level. Age 15 is one of the ages in which less errors are found, although a decrease in quantity of jokes produced is found at this age as well. Therefore, we still cannot associate the u-shape at group 3 to a redescription phase of cognitive development.

The difference between number of errors at one age and another is small. This means that the u-shape that we find in combined errors is still not well established. Had we added other error indicators, the u-shape may have changed. Therefore, the results we have are still preliminary. Moreover, we still cannot explain why the few number of errors made at age 10 are not equaled until age 15.

Baradel and Barbieri (2001) find the lowest point of their u-shape at age 8 while we find it at age 7. This difference could perhaps be partially explained by the little number of subjects we have at age 6 and 7. Another explanation might be that we take into account errors related to the content of the text while they also consider oral performance. Finally, the variable monolingualism

vs. multilingualism could influence results. Despite the differences we find, the use of more subjects from our corpus with a consideration of the language variable could establish whether these differences are definitive.

We asked ourselves about the distribution of languages in the text types we consider and a possible relation between display of multilingualism and errors. Although it has been observed that half of the texts in Spanish from the whole CesCa corpus are found under the label “jokes” (Llauradó, submitted), there is an internal difference inside the label between jokes and other humorous texts regarding language use. Jokes present less Catalan and more code-switching than other texts; Spanish remains very similar. For the case of children that speak other languages at home, the reason for the presence of Spanish is not necessarily that the joke is usually told in Spanish, because other humorous texts that are not jokes present the same percentage of complete texts in Spanish. However, this could be the reason for the presence of more code-switching in jokes than in other humorous texts.

The use of Catalan increases with age, while code-switching decreases. Therefore, although code switching has a function in joke telling, children learn to eliminate it, at least when they perform a written task in a class context. However it is more difficult for them to eliminate code-switching from jokes than from other texts (cf. Tolchinsky and Llauradó, in press).

Code-switching can be free or required by the Joke. It is always free in the other humorous texts. However, since some jokes play on Spanish words, they must be at least partially told in Spanish. Anyway, the code-switching required by the joke is only 1/3 of total code-switching in jokes. Therefore, some code-switching in jokes is free and it might appear for stylistic reasons or because that joke was told in Spanish to that subject.

Translation failures have very few occurrences but, surprisingly, they primarily occur in ages from 8 to 14. Children at ages 6 and 7 successfully tell jokes based on puns. This most likely means they learned the jokes by heart and do not realize the language in which they are telling them.

Free code-switching could be associated with expandable jokes while constrained code-switching could be associated with fixed jokes or rewordable jokes. However, we do not see any parallels between the distribution of language use and types of text used or the errors found.

9. Conclusion

We believe that joke telling is a suitable site for observation of developmental change in the interaction between languages, the relationship between textual structures and their functions, colloquial uses of language, figurative uses of language, control over ambiguity and different shapes of development.

We sought categories of analysis that turned out to be useful in classifying types of response to instructions, types of jokes, language uses and types of errors. However, we believe we must find more precise categories for types of jokes and new categories to add to those we used to code errors.

Although our results are still preliminary, we found a u-shape in quantity of jokes and quality of jokes measured by an index of errors. Despite the u-shapes, overall development was linear and steady when observed by age groups. Different indicators pointed to a difference of the oldest group. More precision in differentiating types of humor and linguistic resources could help us determine whether jokes can be associated to cognitive development in adolescence.

Baradel and Barbieri (2001) also found a u-shape, though at a different age. Another study with more subjects from CesCa corpus and also using L1 Catalan speakers might help clarifying the reason for the differences between our results. Considering other types of errors could also help us establish the qualitative u-shape.

Regarding multilingualism, we found that jokes present less Catalan and more code-switching than other humorous texts. We could also distinguish between free code-switching and code-switching required by the joke. Future studies should use language spoken at home as a variable. The study of joke production in developmental perspective has proven to be a promising domain of inquiry deserving further exploration.

References

- Arts, E. (1988). *Relation of Catalan Jokes with Linguistic and Literary Phenomena*, unpublished doctoral dissertation. Universitat de Barcelona.
- Ashkenazi, O., & Ravid, D. (1998). Children's understanding of linguistic humour: An aspect of metalinguistic awareness. *Cahiers De Psychologie Cognitive*, 17(2), 367.
- Attardo, S., Attardo, D. H., Baltes, P., & Petray, M. J. (1994). The linear organization of jokes: Analysis of two thousand texts. *Humor - International Journal of Humor Research*, 7(1), 27-54. doi:10.1515/humr.1994.7.1.27
- Attardo, S., & Chabanne, J. (1992). Jokes as a text type. *Humor - International Journal of Humor Research*, 5(1-2), 165-176. doi:10.1515/humr.1992.5.1-2.165
- Baradel, C. & Barbieri, M. S. (2001). Ridere e far ridere. *Età Evolutiva*, 70, 98-104.
- Berman, Ruth A. (2004) *Between emergence and mastery: The long developmental route of language acquisition*. Berman, Ruth A. (ed.) *Language Development across Childhood and Adolescence*, Amsterdam-Philadelphia: John Benjamins.
- Bergen, B. & Binsted, K. (2004). *The cognitive linguistics of scalar humor*. Achard, M. & Kemmer, S. *Language, Culture and Mind*, CSLI Publications.
- Brock, A. (2008). *Humor, jokes and irony versus mocking, gossip and black humor*. Antos, G., Ventola, E., & Weber, T. (eds.). *Handbook of interpersonal communication*. Berlin: Mouton de Gruyter. Retrieved from <http://www.loc.gov/catdir/toc/ecip0826/2008037812.html>
- Führ, M. (2001) Some aspects of form and function of humor in adolescence. *Humor - International Journal of Humor Research* 14 (1), 25-36.
- Gershkoff-Stowe, L. & Thelen, E. (2004): U-Shaped Changes in Behavior: A Dynamic Systems Perspective, *Journal of Cognition and Development*, 5:1, 11-36.

- Hetzron, R. (1991). On the structure of punchlines. *Humor - International Journal of Humor Research*, 4(1), 61-108. doi:10.1515/humr.1991.4.1.61
- Horgan, D. (1981). Learning to tell jokes: A case study of metalinguistic abilities. *Journal of Child Language*, 8(1), 217.
- Karmiloff-Smith, A. (1992). *Beyond modularity : A developmental perspective on cognitive science* MIT Press.
- Lefort, B. (1992). Structure of verbal jokes and comprehension in young children. *Humor - International Journal of Humor Research*, 5(1-2), 149-164. doi:10.1515/humr.1992.5.1-2.149
- Llauradó, A., Martí, M. A. & Tolchinsky, L. (Submitted). Written Catalan language throughout compulsory schooling.
- Martin, R. A. (2007). *The psychology of humor :An integrative approach*. Boston etc.: Elsevier.
- Nerlich, B., Todd, Z., & Clarke, D. (1998). The function of polysemous jokes and riddles in semantic development. *Cahiers De Psychologie Cognitive*, 17(2), 343.
- Oring, E. (1989). Between jokes and tales: On the nature of punch lines. *Humor - International Journal of Humor Research*, 2(4), 349-364. doi:10.1515/humr.1989.2.4.349
- Oriol, C. (2000). Jokes about national groups in andorra. *Fabula*, 41(3-4), 285-293.
- Pujolar, J. (1997). *De què vas, tio?* Barcelona: Empúries. [English version: Pujolar, J. (2000). *Gender, heteroglossia and power: A sociolinguistic study of youth culture*. Berlin: Mouton de Gruyter].
- Reyes Pérez, A., Rosso, P., Martí Antonín, M. A., & Taulé Delor, M. (2009). Características y rasgos afectivos del humor: Un estudio de reconocimiento automático del humor en textos escolares en catalán. *Procesamiento Del Lenguaje Natural*.N.43 (Sept.2009).ISSN 1135-5948, Pp.235-243, Retrieved from <http://hdl.handle.net/10045/11718>
- Ritchie, G. D. (2004). *The linguistic analysis of jokes*. London etc.: Routledge.

- Sanders, D. M. (1986). Sign language in the production and appreciation of humor by deaf children. *Sign Language Studies*. 50 (spring), 59-72.
- Schlinger, H. D. (2009). *The Psychological Record*, 2009, 59, 435–448.
- Searle, J. R., & Vanderveken, D. (1985). *Foundations of illocutionary logic*. Cambridge etc.: Cambridge University Press.
- Siegler, R. S. (2004): U-Shaped Interest in U-Shaped Development-and What It Means, *Journal of Cognition and Development*, 5:1, 1-10
- Simarro (2011, December). Los enunciados retroactivos humorísticos como medidores de la competencia léxico-semántica de hablantes de ELE. Paper presented in the Simposio internacional sobre la ironía y el humor, Alacant.
- Tolchinsky, L. (2004) *The nature and scope of later language development*. Berman, Ruth A. (ed.). Language Development across Childhood and Adolescence, Amsterdam-Philadelphia: John Benjamins.
- Tolchinsky, L., & Llauradó, A. (in Press). *Developing a written lexicon in a multilingual environment*. Amsterdam-Philadelphia: John Benjamins.
- Tolchinsky, L., Martí, M. A. & Llauradó, A. (submitted). The growth of the text-embedded lexicon in Catalan from childhood to adolescence.
- Vaid, J. (2006). *Joking across languages: perspective on humor emotion and bilingualism*. Pavlenko, A. (ed.), *Bilingual minds :Emotional experience, expression and representation*. Clevedon etc.: Multilingual Matters.
- Viana, A. (2004). *Acròbates de l'emoció :Exploracions sobre conversa, humor i sentit*. Tarragona: Arola.

APPENDIX

Categories, codes and criteria for codification

Types of response

Category 1

Speech act: Codes the type of response in the adjacency pair instruction-response with a pragmatic perspective. We use speech acts in the sense of Searle and Vanderbecken (1985, p. 53-54).

Code 1 **Directive**

The participant interpreted the instruction as a request to write a funny text.

“un xino i un catala que el xino es cau a laigua i diu gelp gelp gel no pro si bols xampu si” Information about the subject: (age: 6, language: Catalan, Spanish, other, onset: always, id: 549)

Code 2 **Assertive**

The participant interpreted the instruction as a question seeking for information.

“Quan els nens fan tonteries a la classe” (age: 10, language: other, onset: since 1 to 4 years, id: 11622)

Code 3 **Other**

Includes comments on the experimental situation with no other response, change of topic, refusal, not understandable response, etc.

“Re parce io no se riure” (age: 7, language: catalan and other, onset: always, id: 757)

Category 2

Text type: Codes the type of text in the response with a textual perspective. Precisely the previous category and distinguishes two types of directives: jokes and anecdotes/tales; and two types of assertives: more narrative assertives (quasi anecdotes) and less narrative assertives (more descriptive).

Code 1: **Joke**

“en un supermercat can dir per e microfon: senyor perez pugi a la 1ra planta... i al cap d'una estona: senyor perez baixi del test” (age: 8, language: Catalan, other, onset: always, id: 4313)

Code 2: **Anecdote/Tale**

“Un dia hi havi aun park public i havia una balla molt petita i el meu germa anaba a passar per sobre i la Marina va obrir la porta en aquell moment perquè no savia que anaba a saltar, es va donar un petit cop però no li va passar res” (age: 11, language: Catalan, Spanish, other, onset: always, id: 2072)

Code 3: **More narrative assertives**

Quasi anecdote or quasi tale. We assigned this code to texts with two or more dynamic verbs. Dynamic verbs usually correspond to narrative events. This type of assertives usually starts with a particle such as “quan” (when) and refer to an imagined question; the instruction has been interpreted as a question. We do not count static verbs nor the repetition of part of the instruction.

“quan anava corrent que me rrelliscat amb la sorra i el cul a rebut un bon cop” (age: 8, language: Catalan, Spanish, other, onset: always, id: 9170)

Code 4: **Less narrative assertives**

This type of assertives is more descriptive. They could be said to be pointers: they direct the attention of the reader to a situation. We assigned this code to texts with one or less dynamic verbs. They usually start with a particle to refer to the imagined question. Nominal phrases or lists of nouns have been assigned also this code. This type of response could be called an ostensive answer.

“quan s'ha descidat el munade i el pasa port” (age: 8, language: Spanish, other, onset: always, id: 827)

“un xiste” (age: 8, language: Spanish, other, onset: x, id: 772)

Code 5: **Other**

This code has been assigned to other types of text: not understandable texts, texts which are unrelated to the instruction, games that imply teasing, etc.

“- El Pere de * té 5 filles. Es diuen Pata, Peta, Pita, Pota i... - Puta. - No! He dit que el pere de Peneta... O sigui que Peneta és una filla: Pata, Peta, Pita, Pota i Penata.” (age: 11, language: Catalan, Spanish, other, onset: always, id: 364).

Types of Jokes

Category 3

Structure: This category codes the superficial structure of jokes.

Code 1: **Narrative**

This code has been assigned to narrative jokes with or without dialogue. Some jokes present a minimal context set-up. We include also here dialogue with the context implied.

“i avia un espanyol, un portuges, i un xino i van a una cova. I el esp- anyol sen un suroll i samba i va el portuges i tambe san va i el xines sent un suroll i diu el xines que dius a que at compri un v*cini, pos te cinco € para, comprarte un nou. FI” (age: 7; language: Catalan, Spanish, other; onset: more than four years; id: 603)

Code 2: **Riddle**

This code has been assigned to humorous riddles. Riddles are question-answer adjacency pairs.

“quin es el sun duna taula. - que te potes i no pot corre.” (age: 8, language: Spanish, other, onset: more than four, id: 767).

Code 3: **Shortest**

This code has been assigned to the shortest narrative jokes. They usually present one character and one action.

“"Hi va un peix i s'enfonsa". (no en té molta gracia, però em ric molt) xD!” (age: 15, language: Catalan, Spanish, other, onset: more than four, id: 7328).

Category 4

Expandability: this category codes a feature of jokes, the possibility of adding freely comments, expressions or details without compromising the humor.

Code 1: **Expandable**

This code has been assigned to jokes to which it is possible to add details without changing very much the humorous effect. However, there are always limits. According to Oring (1989) one of the Joke-destroying techniques is the surplus-details technique. The attention of the receiver is directed to details that prevent from noticing the humorous device.

“Hi havia una vegada un toro i una vaca el toro salta la valla i li diu a la vaca ola guapa com et dius Blancaneus però millor que em diguis blanca perque les neus estan a la montanya i tu com et dius en dic Gopenis per millor que em diguis ga per

que el penis me le deixan a la valla.” (age: 8, language: Catalan, Spanish, other, onset: always, id: 682).

Code 2: **Rewordable**

This code has been assigned to jokes for which widening is half possible. Only rewording or adding some very few words is allowed. Adding any details before the punchline might make the joke less funny.

“Que li diu Tarzan a un ratoli-que petit i amb bigotis - I tu tant gran i amb panyals!” (age: 10, language: Catalan, Spanish, other, onset: always, id: 927).

Code 3: **Fixed**

This code has been assigned to jokes when adding any details before the punchline might ruin the joke.

“avia un ona Tan Tan Tan Que se convertio en canpana” (age: 8, language: Catalan, Spanish, other, onset: always, id: 11129).

Category 5

Degree: distinguishes first from second degree jokes and different types of second degree jokes (Attardo et al., 1994). Second degree jokes: it is necessary to know a previous joke to understand it or the object of the joke is precisely the usual structure of some jokes.

Code 1 **First degree**

This code has been assigned to jokes that do not play directly on other jokes to create humor

Code 2 **Particular**

This code has been assigned to second degree jokes based on a particular joke shared by the participants in the joke telling session.

“Van 3 i cau el del mig” (age: 10; language: Catalan, Spanish, other; onset: always; id: 1007)

Code 3 **Type**

This code has been assigned to second degree jokes based on a type of joke.

“Va un cargol i derrapa!!! (Quina és la gracia: que no h ha curva!!!)” (age: 10; language: Catalan, Spanish, other; onset: always; id: 4899)

Code 4 **Concept**

This code has been assigned to second degree jokes based on the very concept of joke.

“Ba un xino i es parteix el cul. Ta fet gracia? dons els xinu tampoc.” (age: 8; language: Catalan, other; onset: always; id: 4318)

Category 6

Multimodality: distinguishes the presence of other modes that would appear in an oral performance but can not be transferred to the written text.

Code 1 Verbal

We assign this code when telling the joke in a single mode is possible.

Code 2 Non-verbal

We assign this code when the joke demands gestural support or gestural support is important.

“abia una familia ban en sendra una espelma i el par fa ppp be.” The joke ends when a character puts out a candle with the fingers. (age: 6; language: Spanish, other; onset: from one to four years; id: 4193)

Code 3 Para-verbal

We assign this code when a joke needs a special intonation, voice or singing.

“Juanito va a l'escola i la profe li diu dema eu d'enar disfresats de pastors i diu *a velen castores a velen castores* i la profe el perdona” (age: 6; language: Catalan, Spanish, other; onset: from one to four years; id: 4173). The section in italics must be sung.

Category 7

Comments: the participant writes a response but comments also on it or adds complements like hahaha or emoticones.

Code 1 Comments

“Ara no recordo molt. Però un xiste, que és molt dolent i que le és: Comete el bocadillo de jamón, y jamón se quedó sin bocadillo” (age: 14, language: Catalan, other, onset: always, id: 6264)

Code 2 No Comments

Category 8

Fragmentarity: This category belongs to the description of errors and it refers to the capability to produce a whole text.

Code 1 **No fragment**

We assign this code when a joke starts and ends, although it might lack something in the middle, between beginning and punchline.

Code 2 **Fragment**

We assign this code when a joke starts but does not end.

“va un senyor i veu un alicopter i diu el senyor no es una helicotera” (age: 11, language: Catalan, Spanish, other, onset: always, id: 2102)

Category 9

Completeness: This category codes elements lacking in the middle of the text without counting the punchline.

Code 1 **Complete**

Code 2 **Almost**

We assign this code when a joke is almost complete, that is, when lacks an element that is very helpful but not necessary for the understanding or for the humorous effect.

“És un lladre que es troba a un altre i li diu: - El otro dia me encerraron las salidas del cine. I li diu: - I como saliste? - por las entradas Va respondre.” (age: 11; language: Catalan, Spanish, other; onset: always; id: 2237)

Code 3 **Incomplete**

We assign this code when a necessary element for the humor effect lacks in the middle of the text, especially events.

“Aixo son tres homes NiNgu, ToNto, Naidie Ningu s'acaigut al precipici i NiNgu le esta ajudant ets toNto si molt gust.” (age: 8; language: Catalan, Spanish, other; onset: more than four years; id: 9245)

Code 4 **Take-up**

We assign this code when an element that is taken up later lacks or is not properly presented.

“i avia un espanyol, un portuges, i un xino i van a una cova. I el esp- anyol sen un suroll i samba i va el portuges i tambe san va i el xines sent un suroll i diu el xines que dius a que at compri un v*cini, pos te cinco € para, comprarte un nou. FI” (age: 7; language: Catalan, Spanish, other; onset: more than four years; id: 603)

Code 5 **Character**

We assign this code when a character appears before being presented.

“Un Ingles que estaba en un baixell i el Ingles ba caure a lagua i diu Help i lespanyol diu no tinc help pero si jabon” (age: 8; language: other; onset: always; id: 11242)

Code 6 **Unreferenced**

We assign this code when we do not have a joke of reference to compare it; that is, we recognize it as joke but it could not be documented.

“llavia una vagada 3 tomaques i 1 tomaquet diu quin fret laltre diu tas daguantar pues aguantate tu” (age: 6; language: Catalan, Spanish, other; onset: always; id: 4183)

Code 7 **Too fragmentary**

We assign this code when it is not possible to know how to code the joke because it is too fragmentary. Some longer fragments could lack something.

“habia hum havión que se cHoco y solo havia hum para cahidas y el para cahidas se lo llevo el conductor” (age: 12; language: other; onset: ?; id: 2862)

Category 10

Punchline: This category codes the success or existence of the punchline. Only for jokes that are not fragments have a punchline.

Code 1 **Correct**

We assign this code when punchlines are effective.

Code 2 **Failed**

We assign this code when the punchline is fragmentary or modified in a way that affects heavily the comprehension of the humor.

“an un abio hi ha un frances i una xinu i un aspanyol i al frances diu va fransia i es mata i el xinu diu va xina i al espanyol diu per les escales.” (age: 7; language: Spanish, other; onset: from 1 to 4 years; id: 777).

Code 3 **No punchline**

We assign this code when there is no punchlines. Then, the joke is fragmentary.

“es un nen que fica una xinxeta a la cadirea de la senyoreta i la senyoreta diu avui explicarem la lliso aaa!” (age: 7; language: Catalan, other; onset: always; id: 4268).

Category 11

Repetition: We use this category to code jokes based on repetition as one of the devices to create humor. These jokes create a pattern through repetition that is broken by the punchline. When the pattern is not properly created the punchline cannot contrast (Hetzron, 1991).

Code 1 **Irrelevant**

We assign this code to non-repetition jokes.

Code 2 **same**

We assign this code when a joke teller comments that *the character in the joke does the same* without repeating which are the actions. It is a strategy to avoid repetition in discourse that is not recommendable in repetition jokes.

“Hia un Frances, llitano, arab, i un espanyol, el arab d*iu a tots anem al cementeri i van veuen un fantasme el fantasme diu al espanyol mira la raja de miculo i el espanyol es va corrent i li diu a tots el mateix i es van corent pero queda el gitano i el fantasma li diu mira la raja de mi culo i el g*ita li diu i tu mira la mia” (age: 10; language: Catalan, Spanish, other; onset: always; id: 8834).

Code 3 **Repetition**

We assign this code when there is enough repetition, at least some actions. Enough is 2 times; 2 pulses that establish the patron and then the third discordant pulse.

“Es un Morro un Frances i un Espanyol que es van explicant les seves abentures i diu el Morro: exacanse el pantalo: - Aquesta raja me le feta a Americacyty i diu el Frances ensenyan el baraç: - Pues la meva es de Morrolandiacyty, i salta el Espanyol i esenya la panxa: - Pues això es la Apendicity” (age: 11; language: Catalan, Spanish, other; onset: from 1 to 4 years; id: 6769).

Code 4 **No repetition**

We assign this code when the joke teller does not repeat enough times.

“el Ronaldinho pierde el avion del barsa i tiene que ir con los del madrid. El piloto dice tirar las mochilas que ai muxo peso todos las tiran menos el Ronaldinho. El piloto buelbe a decir agararo al texo que boi a tirar el suelo. El Ronaldinho sice tot el cam i los del madridacen palmas” (age: 8; language: Spanish, other; onset: always; id: 4293).

Category 12

Display of multilingualism: This category codes the different uses of languages in joke telling.

Code 1 **Catalan**

Code 2 **Spanish**

Code 3 **Required**

We assign this code when because the joke requires a section in a particular language. However, when we refer to code-switching in general in the text of our study we mean the sum of all codes that imply code-switching.

“juanito va a l'escola i la profe li diu dema eu d'enar disfresats de pastors i diu *a velen castores a velen castores* i la profe el perdona.” (age: 6; language: Catalan, Spanish, other; onset: always; id: 4173)

Code 4 **Bilingual**

We assign this code when the joke must be obligatory told in two languages (Bilingual joke). Also Humor based on the knowledge of two languages, even if it is only the phonetics or intonation of that language.

“un xino i un catala que el xino es cau a laigua i diu gelp gelp gel no pro si bols xampu si.” (age: 6; language: Catalan, Spanish, other; onset: always; id: 549)

Code 5 **Free**

We assign this code when we observe apparently free Code-switching. When we say free we mean for other functions not directly related to the humor device of the joke

“Eren tres nens que només deien: "nosotros tres nosotros tres", " en bicicleta - en bicicleta, "por el dinero- por el dinero". Es van trobar un assassinat. Un policia va preguntar; Policia - Quien ha sido? Nens - Nosotros tres - nosotros tres Policia- Porque? Nens - por el dinero - por el dinero Policia - a la carcel! Nens- En bicicleta, en bicicleta” (age: 10; language: other; onset: more than four years; id: 429)

Code 6 Translation

We assign this code when there is a translation failure. The joke is not understood because the subject is not aware that the joke requires switching into Spanish and translates everything into Catalan. The result is that the joke is not funny.

“Que li diu el mar als que s'ofeguen? -res.” (age: 14; language: Spanish, other; onset: always; id: 549)

Code 7 No language

We assign this code when the response consists of numbers, a language that cannot be determined or pseudowords.

“lisora les loa le peo” (age: 7; language: Catalan, Spanish, other; onset: always; id: 4218)

Criteria to choose one text when subjects wrote more than one

1. When there is a joke we choose the joke.
2. When both texts are the same type we choose the longer one.
3. When they are equally long we choose the first one written.