

LumaLiDaAudy

Frota, S., M. Vigário, N. Matos, M. Cruz & R. Jordão (2012)

Laboratório de Fonética, CLUL/FLUL

Transcription Criteria

1. Child Utterance
2. 'Phonetic' Orthography of child utterance
3. Target Utterance
4. Target Orthography
5. Meaning
6. Prosodic Phrasing
7. Observations on other fields of the database

1. Child Utterance

General criteria

Phonetic transcription was made in IPA. Three researchers independently transcribed the child speech and the phonetic transcription was validated by a fourth researcher.

- (a) Word stress is marked immediately before the syllable.
- (b) Words (corresponding to orthographic words) are separated by a blank space, except in the case of pronominal enclitics.
- (c) All prosodic words are marked with a primary word stress.
- (d) The segments usually taken to be labialized in the target – [kw] and [gw] – were always transcribed as consonant plus glide (Cw) in child speech: e.g. ['kwatu], target ['kwartu].
- (e) The diacritic `:' is used after a lengthened segment in its production.
- (f) When the same sequence is uttered several times, the entry is repeated the exact number of times the sequence was produced.
- (g) Short pauses are indicated with `/' and long pauses with `//'.

2. 'Phonetic' Orthography of child utterance

2.1. Why a 'phonetic' orthography

Following LumaLiDaOn (Frota, Vigário & Jordão 2010) criteria (<http://www.fl.ul.pt/LaboratorioFonetica/res/LumaLiDaOnTranscriptionCriteria.pdf>), this field was introduced to enable the tool FreP (www.fl.ul.pt/laboratoriofonetica/FreP) to run on child data, so it is of special interest both to FreP users and to researchers that want to take advantage of the facilities provided by FreP to get the frequency of phonological objects and patterns out of child speech data.

2.2. General criteria

- (a) The orthography includes only the segments that are present in the phonetic transcription, i.e. the number of segments in both transcriptions matches.
- (b) Coalescence is respected unless it is independently known that the child is able to produce all the elements involved in the process.
- (c) Orthographic marks of stress are inserted to ensure that the location of stress in speech is computed correctly. Thus these marks signal stress position and not vowel quality.
- (d) Non-full vowels in the phonetic transcription are included in the orthography.
- (e) Deleted vowels are excluded from the orthography, except in the case of deleted [i]: e.g. [k 'tẽ] > que tem.
- (g) The insertion of the [i] sound is transcribed as 'e' in the orthography: e.g. ['papwi] > papue.
- (i) Target names are written with an initial capital letter.

3. Target Utterance

3.1. On the nature of the phonetic transcription of the target

The phonetic transcription of the target is generally conservative in the sense that it closely follows the orthography when more than one option is available in the language. By and large, the transcription reflects the lexical phonology. External sandhi phenomena are included only when obligatory or very frequent.

3.2. General criteria

- (a) On the lexicon: the target lexicon is indicated and not specific forms used by the child (e.g. child 'nana' > amarelo; child 'pepé' > sapato).
- (b) If clear from the child utterance and context, grammatical information like number and gender are included.
- (c) Clitic words are also included.
- (d) The word order of the language is followed.
- (e) Word initial vowels are transcribed following the orthography (e.g. [eli'fɛti] 'elefante', [obri'gadɐ] 'obrigada').
- (f) Other cases with possible alternative productions are also transcribed following the orthography (e.g. 'piscina' with two [i] vowels and 'crescido' with one [ə] and one [i] vowel).
- (g) The segments usually taken to be labialized in the target – [kw] and [gw] – are transcribed as such.
- (h) Voiced plosives are not transcribed as non-strident fricatives in between vowel position.
- (i) The 'ing' sound is transcribed as a velar nasal ([ŋ]).
- (j) Ambisyllabicity is not marked.
- (k) All prosodic words are marked with a primary word stress.
- (l) The homophonous forms of the indefinite article and the numeral ('um', 'uma') are unstressed in the case of the article and stressed in the case of the quantifier and in case of doubt between one and the other.

4. Target Orthography

The target orthography is simply the orthographic counterpart of the target phonetic transcription. Proper names are given with an initial capital letter.

5. Meaning

This field is intended to make clear what the exact meaning of the child utterance is. It thus may correspond to the target orthography or include additional elements explaining the intended meaning of the utterance.

6. Prosodic Phrasing

Labelling for prosodic phrasing followed the general properties of the prosody of Portuguese (Frota 2000, 2014; Vigário 2003), together with information on the specific features of Luma's speech as described in Frota & Vigário (2008), Frota & Matos (2009).

As a rule, we opted for minimal phrasing, and thus the prosodic phrasing deviates from adult prosodic structure in three ways:

(1) optional phrasing effects were not considered (e.g., weight effects and compound prosodic domains);

(2) a PW was usually mapped into PhP, following the tendency described in Frota & Vigário (2008) and Frota & Matos (2009);

(3) prosodic levels between PW and PhP were not considered, beyond the assumption that the unstressed syllables of clitics adjoin to the adjacent PW.

7. Observations on other fields of the database

The Context field includes information on the context of production (such as previous and/or following discourse, description of the scenario, or description of actions).

The Comment field comprises all sorts of comments relevant to the nature of the phonetic transcription and/or the meaning of the utterance produced. Paralinguistic information is also annotated here, such as scream, cry, laugh, cough.