

# **SYNTAX, LEXICON, MORPHOLOGY AND PRAGMATICS IN AN EARLY STAGE OF BILINGUAL ACQUISITION: A COGNITIVE APPROACH**

HANNA JAKUBOWICZ BATORÉO  
(Universidade Aberta, Lisboa / CLUNL)

*ABSTRACT: In the present study we discuss data relating to bilingual acquisition of Polish and European Portuguese by a Polish-Portuguese child. The term 'bilingual acquisition' is used to refer to a child's regular simultaneous exposure to two languages after birth and during the first years of life (cf. De Houwer, 1990, 1995; Lanza, 1997; Deuchar and Quay, 2000). The report is based on longitudinal data samples of the girl's spontaneous language production in her first years (cf. Batoréo, 2007), and especially on her three-staged developmental sequence of the mixed period of language acquisition. Special attention is drawn to the mixing patterns observed in nominal and verbal constructions, in which part of the linguistic information is brought from one input language and another part from the other tongue. The notion of 'interface' is questioned and a cognitive (usage-based) approach to language acquisition is adopted (Langacker, 2009; cf. Rowland and Theakston, 2009). The syntax, lexicon, morphology and pragmatics of the early bilingual acquisition are discussed not as modular or interface phenomena but as multiword strings analysed on the basis of a holistic usage-based approach.*

*KEYWORDS: bilingual acquisition; mixed period of language acquisition; cognitive (usage-based) approach to language acquisition; acquisition of Portuguese; acquisition of Polish*

## **1. Cognitive versus modular approach to language acquisition**

Language acquisition has been largely investigated in terms of linguistic theories, as for instance Generative Grammar, that argue for the existence of so called separate language "modules", and of interfaces covering the space existing among them. Thus, the process of acquiring a language, according to Optimality Theory (cf. Freitas, 2003, among others), implies the acquisition of the ranking of constraints and following learning algorithms until the acquisition of the target system. According to these theoretical frameworks,

the ultimate aim of this sort of investigation is to determine the architecture behind grammatical processing:

The authors use different theoretical frameworks and formal tools to evaluate evidence from the perception and the production modules. Their main goal is however the same: they all seek to a better understanding of the path of language development as means to access the architecture behind grammatical processing.

Freitas (2003: 5)

Nevertheless, there are other theoretical backgrounds that are cognitively oriented, which means that they are not interested in abstracting architectural modules, as they rely upon abstracting of linguistic patterns from usage events. The concern of this sort of investigation is to answer the following research questions: “*How are linguistic patterns abstracted from usage events? What precisely is abstracted? In what form is it stored or represented? How is it then used?*” (Langacker, 2009: 628). Thus, the concern is with a dynamic conception that recognises the temporal dimension of linguistic structure as an aspect of cognitive processing:

These questions presuppose a particular theoretical orientation, defined by several organizing assumptions. For one thing, language is seen as an integral part of cognition, not a separate “module” (hence cognitive linguistics). Moreover, the pivotal factor is meaning, rather than abstract form (hence the symbolic view of grammar). And from these considerations it follows that language is learned through meaningful use, rather than being innate (hence the usage-based approach).

Langacker (2009: 628)

According to the cognitive (usage-based) approach to language acquisition (cf. Rowland & Theakston, 2009), children learn multiword strings and not just single words for their subsequent assembly, supporting an account of

language development based on learning pieces of language from the input mapped to child-based meanings, with the development of a more schematic and abstract inventory of conventionalised constructions.

Lieven, Salomo & Tomasello (2009: 505)

Departing from previous studies by Batoréo (1989, 1991, 1998), the analysis that we present in this paper is based on longitudinal data samples of a Polish-Portuguese bilingual girl’s spontaneous (usage-based) language production in her first years of life (cf. Batoréo, 2007, 2011), and based especially on her three-staged developmental sequence of the mixed period of language acquisition. Special attention is drawn to the mixing patterns observed in the case of constructions with nouns and verbs as their nuclei, in which part of the linguistic information is brought from one input language and another part from the other.

The analysis shows that what in one language is realised as morphology can be produced in the other as e. g. morphosyntax, syntax or pragmatics. These results imply that studying different ‘modules’ in a modular (and interface) theoretical frame is misleading, as a given module may account for different language material in different particular languages.

## **2. Bilingual acquisition**

The main aim of case studies in bilingual acquisition is to explore its implications for linguistic theory, some of them for language acquisition theory in general, and/or for theories of bilingual acquisition in particular. Specific implications for bilingual acquisition include the questions of (i) whether originally a bilingual child had one or two linguistic systems, (ii) what criteria should be used in identifying one versus two systems, and (iii) what the most important determinants of language choice are for the developing bilingual.

The revision of bilingual acquisition literature (see: Cruz-Ferreira, 2003, cf. chap. 1. ‘*Lingualism*’ and *bilingualism* 2003: 45-49) shows that the definition of a child as bilingual suffers from the same indeterminacy as the definition of any speaker as bilingual. In her thorough revision of the (English) literature on the subject, the author points at monolingual theorisation as the main reason for all the misunderstandings of bilingual acquisition centred research:

It is clear that as much insight into bilingualism can be gained from monolingual-based theorisation as to siblinghood from within a framework designed to characterise a single child. (...) The consensus in research on bilingualism seems then to be to approach the use of two languages from the perspective of one of them or of a merged version of both. (...) If we take early bilingual speech as an instance of the use of language and not of the use of particular languages, we may open the way to quite different conclusions about bilingualism. Child systems are systems in the making, and we are therefore dealing with the process of acquisition, not its product.

Cruz-Ferreira (2003: 46-47)

The term ‘bilingual acquisition’ is used in the present study to refer to a child’s regular exposure to two languages after birth and during the first years of life (cf. De Houwer, 1990, 1995; Lanza, 1997; Deuchar & Quay, 2000).

## **3. Bilingual acquisition process as system-in-the-making: Polish-Portuguese data analysis**

The data in our bilingual acquisition corpus come from Marta, who was born in Lisbon, in a family of a Polish mother and a Portuguese father, with

European Portuguese being used at home as the family language. From the very beginning the 'one-person-one-language' strategy was adopted in language interaction as a guarantee of desired bilingual acquisition (cf. Fantini, 1985; Saunders, 1982, 1988). Polish was used only by the child's mother and European Portuguese by her Portuguese father, family, friends and all other interlocutors in every-day life.

This strategy proved to be quite fruitful in the first two and a half years of Marta's life when she was in her mother's care. At 2;06, the child joined a Portuguese kindergarten. At that time her possibilities of verbal interaction in Polish started to decrease, while the diversity of interaction possibilities in Portuguese experienced a remarkable growth. Due to the enriched Portuguese environment and the growth of differentiated Portuguese language interactions with child carers, other adults and new friends, the strategy used successfully in the first years began to fail, as Marta's mother started using Portuguese while addressing her daughter in monolingual Portuguese group interactions. Thus the presence of the Polish input, non-balanced from the very beginning and practically restricted to the mother's language production, started to decrease, becoming partly substituted by mixed bilingual input. Marta's language data prove that she can be considered a balanced Polish-Portuguese bilingual until the age of 2;6, when she began to lose her balanced capacities. At age 5;0 she was considered a non-balanced bilingual dominant in Portuguese, and eventually a Portuguese monolingual with some (uneven) command of Polish (cf. Batoréo, 1989, 1991, 1998)

A great number of studies of bilingual acquisition confront and discuss the problem of the existence of one or two linguistic systems in a bilingual child:

Studies in bilingual acquisition have been dominated by the question of whether children acquiring two languages simultaneously start out with one linguistic system that later develops in two, or whether they have two systems from the very beginning. Klausen, Subritzky and Hayashi (1993: 63) label these the 'one-system-model' and the 'two-system-model' respectively. De Houwer (1995: 231-5) provides an overview of the one-system model and its critics, but also asks whether it is appropriate to raise this question in relation to developing bilinguals under the age of 2 years.

Deuchar & Quay (2000: 111)

Our data are decisive in showing that in the process of her bilingual acquisition the first four years of Marta's life can be considered in terms of a three-stage developmental sequence (cf. Batoréo, 1998). The first part of the sequence can be seen as a mixed period in language acquisition, subdivided subsequently in stages 1 and 2, and followed by a post-mixed period (stage 3), as some examples of language production illustrate below, in subsequent sections 3.1 and 3.2.:

(i) The 1st stage (age 0 – 2;2) can be considered a 'one language system' when the child language production cannot be classified as either Por-

tuguese or Polish, but rather as a ‘Luso-Polish language’, with some mixed utterances and mixed word forms made up of both Portuguese and Polish morphemes;

(ii) The 2nd stage (age 2;2 – 3;6) can be considered a transitory period, when a progressive bifurcation of one mixed language system, i.e., the ‘Luso-Polish language’, into two separate languages (Polish and Portuguese) takes place. At 2;2 there begins a ‘self-code-switching’ phenomenon, when the child produces utterances in one language and repeats them immediately in the other without mixing, but juxtaposing them instead as if they were some ‘self-translation’ products. Then, at age 2;4, the earliest systemic code switching appears as a function of the Participant Variable. As the mixed constructions are still used and the separation of the two systems is in progress, the utterances that are effectively produced belong to three (rather than two) types: Portuguese, Polish and mixed ‘Luso-Polish’;

(iii) The 3rd stage (after age 3;6) corresponds to the post-mixed period in language acquisition. That is when a new type of translation in a given language setting arises – as a function of the well defined Addressee variable – while self-translation ceases. Although metalinguistic awareness can be detected from the very early age of language production, a marked growth of metalinguistic awareness becomes overt in this period with code-switching as a function of the categories Participant, Topic and Setting

In order to illustrate the three-stage developmental sequence mentioned above we shall focus on two systems-in-the-making: verbal (section 3.1) and nominal (section 3.2.). It is important to take into consideration that both European Portuguese and Polish are Indo-European languages of rich verbal morphology, but only Polish shows very rich nominal morphology (being nominal categories – nouns, adjectives, different classes of pronouns – marked for seven cases), whereas European Portuguese maintains case marking only on personal pronouns and not on nouns. Polish children acquire the morphology markers of their native tongue as early as age 1;6 – 2;0 (cf. Weist *et al.*, 1984).

In order to make the reading of the examples easier, in the examples presented below, the **Polish** morphemes are marked in bold and the *European Portuguese* ones are marked in italics.

### 3.1. The verbal system-in-the-making

In the mixed period, the child uses mixed Polish-Portuguese utterances in which individual verbs and verb periphrastic constructions show up as constructs of mixed lexical and grammatical morphemes of the two languages to which the child is exposed (see examples (1), (2) and (3) below):

(1)a (at 2;2)

*deixa (a) pani trabalhar*  
 let (the) woman/lady work  
 'Let the woman work.'

(1)b (at 2;4)

*deixa p(rz)ytuli(ć)*  
 let (me) hug  
 'Give me a hug.'

(2) (at 2;3)

*vamos dmuchar*  
 let's blow (the nose)  
 'We shall blow the nose.' / 'Let's blow the nose.'

(3)a (at 3;2)

*może lavać głowę?*  
 may she wash head ACC?  
 'May / can I wash my hair?'

(3)b (at 3;2)

*jà umyłaś głowę?*  
 already have you washed head ACC?  
 'Have you already washed your hair?'

Since the age 2;2, mixed Polish-Portuguese constructions have been observed in the child's language system. These constructions tend to be of one language (morpho)syntax and another language lexicon, as shown in examples (1 a, b) and (2), but the linguistic choice in the mixed construction is not linear at all.

In example (1), at the same age and in the same sort of periphrastic Portuguese structure 'let sb. + Infinitive', Marta shows that she has two different production options: she uses either the Portuguese infinitive (example (1a) or the Polish one (example (1b)). The infinitive forms '*trabalhar*' ('to work'), in Portuguese, and '*przyculić*' ('to give a hug'), in Polish, are constructed by two different morphemes: the lexical one '*trabalh-*' and '*przycul-*' and the infinitive grammatical marker '*-ar*', in Portuguese, and '*-ić*', in Polish.

At the same time, on the other hand, in example (2), in a future two-verb construction 'ir + Infinitive' used in both languages, the girl chooses to form a 'Luso-Polish infinitive', made up of a Polish lexeme '*dmuch-*'

from the verb *dmuchać* ('to blow') and a Portuguese infinitive marker *-ar*' (see also example (3a)).

In examples (3a) and (3b), produced one year later, at age 3;2, Marta shows that she can refer to the same interaction event with mixed language material, processing the mixing activity in two different ways, cross-using not only lexical and grammatical morphemes from both languages but also using pragmatic strategies – as described below – acceptable in one of them and not acceptable in the other.

Thus, in (3a), the whole utterance seems to be Polish, with only one Portuguese lexeme *'lav-'* from *'lavar'* ('to wash') attached to the Polish infinitive marker *'-ać'*. Nevertheless, the whole utterance is pragmatically Portuguese, as the child – asking her mother for the permission to wash her hair (= head) – uses the self-reference in the 3rd person singular, the form frequently used in standard EP child reference system but unacceptable in its Polish counterpart. Thus, the grammatical form used is *'może'* ('may/can she?') instead of expected Polish *'mogę'* ('may/can I?'). The expression corresponds to *'lavar a cabeça'* / *'umyć głowę'*, with Accusative marking of the noun in Polish, which means literally 'to wash one's head' (in the meaning of 'to wash one's hair'). On the other hand, in (3b), in an analogous context, the morpheme crossing does not occur and the whole utterance seems to be Polish, yet its perfectivity is reinforced by an extra perfective Portuguese marker *'já'*, equivalent to *'already'*, strategy not necessary in Polish in this context. The reinforcement shows the necessity of using the Portuguese strong perfective overt marking, which may be the evidence of not distinguishing the perfectivity inherent in the Polish verb *'umyć'* ('to wash' Perfective) as contrasted with its Imperfective counterpart *'myć'* ('to wash' Imperfective).

The analysis of the verbal system examples presented above shows that, if at the first glance the child may arguably tend to develop her mixed language making it up with Polish lexicon and Portuguese grammar, the evidence seems inconclusive. At this stage, the girl knows how to form infinitives with lexical and grammatical morphemes, forming correct forms in both languages, as well as mixed Luso-Polish forms, but she does not know how to use pragmatic strategies that differ from one language to the other.

### 3.2. The nominal system-in-the-making

Like the case of the verbal constructions presented above, the nominal system is constructed by mixing the material of both languages, as can be observed in examples (4a, b) and (5a, b) below.

(4)a (at 2;3)

chlebek    *com*    mas(1)o  
 bread     with    butter (Nom)  
 'Bread and butter.'

(4)b (at 3;0)

**chlebek**    **z**    **maselkiem**                    *fechado*  
 bread     with butter (diminutive) (Instr) closed (= a sandwich)  
 'A bread and butter sandwich'

(5)a (at 2;2)

**Nie ma**                    **wagi.**  
 there is/ are no        scales (Gen.)  
 'There are no bathroom scales.'

(5)b (at 2;5)

Marta            nie ma,        mamusia,    Marta            nie    ma.  
 Marta(Nom)    is not here, Mummy,    Marta (Nom)    is not here.  
 'Marta is not here = you cannot find me'

As can be observed, the nominal constructions illustrated above in examples (4) and (5) are made up with Portuguese and Polish morphemes, both lexical and grammatical.

In examples (4a, b), the child refers to 'bread and butter'. In the first case (4a), at age 2;3, she uses Polish lexical items 'chleb' ('bread') and 'maslo' ('butter') and joins them with a Portuguese connector 'com' ('with'). The resulting construction 'chleb com maslo' does not resemble the corresponding Polish normative construction 'chleb z maslem', as the second lexical item is not marked in the Instrumental case as expected. At that moment, the child uses an 'unmarked (Nominative?) + com + unmarked (Nominative?)' schema, instead of a 'Nominative + com + Instrumental', as required in Polish.

Nearly a year later, at the age 3;0, the child is already able to produce the correct corresponding construction in Polish (using a diminutive form 'maselko' – 'little bread') with the correct case marking (4b). The interesting aspect of the example is that the girl uses a pragmatic strategy explaining that she wants a 'bread and butter sandwich', but as she does not know the word 'sandwich' she explains that it must be 'fechado' ('closed') in Portuguese. Here both the lexicon and grammar are Polish but the pragmatic strategy is used in Portuguese.

At the age of 2,2 and 2;5 the dominance of case marking system cannot be taken for granted: in the same construction 'There is no + Noun in Geni-

tive', the child uses the noun both marked and unmarked, as in '*nie ma wagi*' ('there is no scales GEN') (example 5a), and unmarked as in '*nie ma Marta*' ('there is no Marta NOM', instead of correct '*nie ma Marty* GEN', in example 5b). The examples show that the correct Polish case marking is in development and is not produced in a more consistent way until approximately age 3.

Whereas both Polish and Portuguese have obligatory grammatical (inflectional) markers in the verbal system, providing transparent 'built up' examples, with grammatical markers coming from both systems, as observed above in section 3.1., in the case of the nominal production we face a grammatically rich system of case markers in Polish and no case marking in nouns at all (marker 0) in contemporary European Portuguese (with the exception of the fixed case marked personal pronoun system). This explains why the language data examined display Polish noun lexemes used with (correct) case markers (examples (4b) & (5a)) next to (incorrect) usage of unmarked forms (examples (4a) & (5b)), as if the global grammatical system of the child was Portuguese at the morphosyntactic and syntactic levels and the lexical material inserted in it were of Polish origin, not allowing grammatical case marking (marker 0). This analysis shows that Marta's data are very different from the Polish data of monolingual children (as stated above in introduction to section 3), who acquire their flexional morphology even before they are two.

#### 4. Discussion and conclusions

The bilingual Polish-Portuguese corpus data show that from the age of 2;2 to 3;6 the child uses only one language in development that is mixed in character that can be labelled as 'Luso-Polish'. It is composed of (i) Polish, (ii) European Portuguese and (iii) mixed language where syntactic constructions of one or the other language (or corresponding to both of them, when they coincide) appear "filled up" with morphological and lexical material from both languages. In these construction there is a certain tendency to use grammar from the "stronger" child's language (European Portuguese) and fill it in with the "weaker" (Polish) lexicon, whereas pragmatic strategies happen to be predominantly Portuguese. Nevertheless, this apparent learning schema is not a very regular one, as morphological richness and grammatical architecture vary between the two languages. The examples show that what is pure syntax in one input language happens to be a complex morphosyntactic construction in the other or that lexical material can be fundamental to pragmatic strategies (cf. description of examples (4 a and b)); they show us also that all this material leads the child to develop a merged system of a mixed language where the frontiers are fuzzy, blur easily, and do not allow to distinguish clearly what can be called a "module"

and what an “interface” or where one so called “module” begins and the other ends, as can be observed above in examples (3a) and (4a & b).

The mixed period evidence from bilingual language acquisition makes us claim that a bilingual child learns pieces of language from the input belonging simultaneously to different so called “modules” and this input is mapped to child-based meanings, with the development of a more schematic and abstract inventory of conventionalised constructions.

It is worth mentioning that different tongues with different morphological characteristics taken into consideration in bilingual acquisition can bring different results as far as the age of language acquisition is considered. Thus, for example, when compared with the language data from Marta’s corpus, other data of bilingual child acquisition can lead us to different results, as exemplified in the case of the English-Spanish speaking child reported on by Deuchar and Quay (2000). The visibility of morphosyntactic merging seems much stronger in our case and lasts longer in the child development (till 3;6), probably due to language dependent phenomena, such as a rich overt case marking in the Polish nominal system (versus no case marking on nouns in Portuguese, Spanish or English). Nevertheless, bilingual child acquisition appears to follow the same overall pattern of language development, with mixed and post-mixed stages, regardless of the particular languages taken into consideration, even if they occur in some children earlier than in others:

Thus we have seen that, in our data, utterances can be classified as either English or Spanish by about age 1;11, on the basis of language specific morphology. We were then able to go and identify language specific syntax, or two systems, in the utterances we analyzed from ages 1:11 to 2;3.

Deuchar & Quay (2000: 87)

The data presented in the present paper let us corroborate the premise forwarded in the very beginning and defended by the cognitivists (cf. Langacker, 2009) that the *pivotal factor in learning is meaning*, rather than abstract architectural form (cf. Freitas, 2003), which means that *language is learned mainly through meaningful use*, rather than being *predominantly determined by universal architectural abstract form composed by grammatical modules and interfaces*.

## References

- Batoréo, H. J. (1989). *A Categoria Linguística Aspecto no Discurso Conversacional de uma Criança Bilingue aos Cinco Anos de Idade*, Dissertação de Mestrado em Linguística Portuguesa Descritiva, Faculdade de Letras da Universidade de Lisboa, Lisboa.
- Batoréo, H. J. (1991). *Crescer Com Duas Línguas: Aquisição da Linguagem por uma Criança Luso-Polaca até aos Cinco Anos de Idade*, ms.

- Batoréo, H. J. (1998). Some Morphosyntactic Phenomena in the Mixed Period of Language Acquisition in a Polish-Portuguese Bilingual Child: Acquisition of Case Markers. *Psychology of Language and Communication* 2(1), pp. 63-73.
- Batoréo, H. J. (2007). Bilingual Acquisition Revisited. Some Implications of a Polish-Portuguese case study twenty years later. In *Textos seleccionados da Conferência on Iberian and Slavonic Cultures Contact and Comparison*. Lisboa: Faculdade de Letras da Universidade de Lisboa, pp. 321-333.
- Batoréo, H. J. (2011). Aquisição Bilingue: Teorias e Aplicações. In M. H. M. Mateus *et al.* (ed.) *Bilinguismo e Educação Bilingue*, Colóquio, Fundação Calouste Gulbenkian, Iltec, 20-21 June, 2011.
- Cruz-Ferreira, M. (2003). Two prosodies, two languages: infant bilingual strategies in Portuguese and in Swedish. *Journal of Portuguese Linguistics* 2(1), pp. 45-60.
- De Houver, A. (1990). *The Acquisition of Two Languages from Birth: A Case Study*. Cambridge: Cambridge University Press.
- De Houver, A. (1995). Bilingual Language Acquisition. In P. Fletcher & B. MacWhinney (eds.) *The Handbook of Child language*. Oxford: Blackwell, pp. 219-250.
- Deuchar, M. & S. Quay (2000). *Bilingual Acquisition. Theoretical Implications of a Case Study*. Oxford: Oxford University Press.
- Fantini, A. E. (1985). *Language Acquisition of a Bilingual Child: A Sociolinguistic Perspective (to Age Ten)*. Clevedon: Multilingual Matters.
- Freitas, M. J. (ed.) (2003). Special Issue of *Journal of Portuguese Linguistics: On the Acquisition of Portuguese*. Lisboa: Edições Colibri – AEJPL.
- Langacker, R. W. (2009). A dynamic view of usage and language acquisition. *Cognitive Linguistics* 20(3), pp. 627-640.
- Lanza, E. (1997). *Language Mixing in Infant Bilingualism: A Sociolinguistic Perspective*. Oxford: Clarendon Press.
- Lieven, E., D. Salomo & M. Tomasello (2009). Two-year-old children's production of multiword utterances: A usage-based model. *Cognitive Linguistics* 20(3), pp. 481-507.
- Mateus, M. H. M. *et al.* (ed.) (2011). *Bilinguismo e Educação Bilingue*, Colóquio, Fundação Calouste Gulbenkian, Iltec, 20-21 June, 2011.
- Meisel, J. M. (ed.) (1994). *Bilingual First Language Acquisition: French and German Grammatical Development*. Amsterdam: John Benjamins Publishing Co.
- Romaine, S. (1995). *Bilingualism*. 2<sup>nd</sup> edition, Oxford: Blackwell.
- Ronjat, J. (1913). *Le Développement du langage observé chez un enfant bilingue*. Paris: Librairie Ancienne H. Champion.
- Rowland, C. & A. Theakston (eds.) (2009). Special issue of *Cognitive Linguistics: Cognitive approaches to language acquisition*. Berlin / New York: Mouton de Gruyter.
- Saunders, G. (1982). *Bilingual Children: Guidance for the Family*. Clevedon: Multilingual Matters.
- Saunders, G. (1988). *Bilingual Children: From Birth to Teens*. Clevedon: Multilingual Matters.

Weist, R. M., H. Wysocka, K. Witkowska-Stadnik, E. Buczowska & E. Konieczna (1984). The defective tense hypothesis: On the emergence of tense and aspect in child Polish. *Journal of Child Language* 11, pp. 347-374.