

# Jadwiga Cook

---

## Working with Bilingual Children. Remarks on the Methodology of Recording and Transcribing Children's Speech

---

Lingwistyka Stosowana / Applied Linguistics / Angewandte Linguistik nr 17,  
15-28

---

2016

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej [bazhum.muzhp.pl](http://bazhum.muzhp.pl), gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

**Jadwiga COOK**

University of Wrocław

## **Working with Bilingual Children – Remarks on the Methodology of Recording and Transcribing Children’s Speech**

### **Abstract:**

The analysis constitutes a part of a larger research project concerning the expression of spatial relations and movement in Polish-French and Polish-English bilingual children’s speech (focusing especially on the signs of interference or code-mixing). Since 2004, when Poland entered the European Union, the number of mixed couples, and therefore the number of bilingual children living in Poland, has increased significantly. However, the research concerning this group of young bilinguals is up to this point almost non-existent. The focus of the study are the methodological aspects of gathering research material among Polish-English and Polish-French bilingual children aged from 4 to 7 living in Poland. Its aim will be to describe and answer the methodological problems arising while working with this group of bilingual children. The aspects described in the paper will be: 1) the speech data collection, especially the process and criteria of the speakers’ selection and conducting the recordings with small children; 2) transcribing English, French and Polish children’s speech into the CHAT format.

### **Introduction**

The opening of borders after joining by Poland the European Union in 2004 undoubtedly widened the mobility possibilities for the Polish people and it increased the number of foreigners coming to work and live in Poland. We cannot call Poland a multicultural land, but we do observe more and more couples of mixed nationality and language, and more bilingual children in the playgrounds, pre-schools and schools. A linguistically mixed family in Poland most often involves a Polish mother and a foreign father.

The phenomenon of child bilingualism related to the Polish territory is relatively new and so far has been rarely researched. We can list a few publications about Polish bilinguals, most of which concern Poles (children or adults) living abroad and the Polish language being the heritage language (e.g. R. Laskowski 2009, M. Mróz 2011, M. Błasiak 2011, U. Paprocka-Piotrowska, G. Komur-Thillooy 2012). The publications about the language of bilinguals living in Poland are scarce, too. Only a master’s thesis prepared at the Jagellonian University in Kraków (K. Košťálova 2010) can be mentioned in this regard.

Our research project focuses on Polish-English and Polish-French bilingual children living in Poland and as such it will hopefully become a step towards filling this gap. The aim of the project is to describe and analyse how bilingual children talk about space and movement. The research has, as a starting point, the work of M. Hickmann

(2012), who investigated the field among monolingual English, French and German speaking children, coming to the conclusion that the development of syntactical structures concerning space and movement depends on the one hand on the child's age and, on the other hand, on the language the child speaks. We are interested in how the same type of structures function in the speech of bilingual children. The objective of our project is to compare how children with English and French as the second language acquired simultaneously describe spatial relations and movement. We will be looking for the similarities and differences in this specific language behaviour in two groups of bilingual children, the variable being the type of the language Alpha. This approach would make the research cross-sectional in design.

The present paper is an introductory part of the project and it focuses on some methodological aspects of working with bilingual children: the search for and selection of the participants, data collection (audio recordings) and data transcription. The last point seems to be of particular importance, as the development of transcripts, the question of what is transcribed and how, has so far received little attention in research literature, as C. Davidson (2010: 115–116) points out.

## **1. Search for and selection of participants**

### **1.1. The definition of bilingualism**

Before beginning to look for participants for the study, first it had to be decided what the term *bilingual* means in the context of the present research. The term has already acquired many definitions, some of them vague or mutually contradictory and taking under consideration different aspects of the phenomenon. As C. Hoffmann (1991: 31) remarks, “bilingualism is a relative concept” and there is no definition that would be “equally acceptable, or useful, for everybody approaching the subject”. Therefore, determining the meaning of the term *bilingual child* for the use of the project was crucial. Bilingualism is considered in a narrow sense of the term. A bilingual child will be a child who acquires two languages simultaneously, from birth, in a natural context and manner, and who can communicate in those two languages on a daily basis with a similar level of competence. Their bilingualism would then be natural, early, simultaneous and is hoped, relatively balanced. Having adopted Jürgen Meisel's term *Bilingual First Language Acquisition*, we are treating both languages as simultaneously acquired. Hence, we did not want to treat them as *first* and *second* or *mother tongue* and *foreign language*. To make the distinction between them, we follow the terminology by A. de Houwer (2009: 2) – language A (Polish, in the case of this study) and the Alpha language (English or French).

### **1.2. Criteria for selection**

Having adopted this approach to the term, we were able to list the selection criteria for the participants. The children who would take part in the research have to:

- be a member of a Polish-British or Polish-French family with a Polish mother and a British/French father;
- be born in Poland or have moved to Poland in the first 6 months of their life;
- have had regular input in both languages from birth;

- present no speech delay or dysfunction;
- be, at the time of the recording, between 4 and 7 years of age.

The lower limit is drawn at 4 years old, based on the fact that at the time children are capable of constructing autonomous utterances, without an adult person's help (U. Paprocka-Piotrowska/ G. Komur-Thillooy 2012: 133). Also the ability to talk about space and localisation becomes more „sophisticated”, children are able to describe the objects they see using locating vocabulary. The upper limit of 7 years is set because it is at this age that children begin their education at school (although this is now in the process of changing to 6) and all start to read and write in Polish, which makes the position of this language even stronger.

### 1.3. The search for candidates

To find the participants, we first asked amongst our family and friends and in this way a few bilingual children in our city were found. We also posted a question on the internet forum for young mothers that we have been a member of, and thereby found another few families from other cities in Poland. Occasionally we came across a bilingual family by coincidence, in which case we came forward and briefly explained what the project is about, what the recordings were like and asked the parents if they would be willing to participate.

Unfortunately, it has not been possible to include in the project all of the children identified. One factor which made us remove some of the names from the list was the fact that not all of the children met the criteria. Some were bilingual, but both parents were Polish, the father, a teacher of English who has decided to speak to them in English. Other problems resulted from the parent's attitude towards their children's bilingualism. In some other cases the foreign father spoke Polish well and did not insist on speaking to his children in his own language and, with time, the whole family switched to communication in Polish. In one case, the father of a 5-year-old girl said that he was beginning then to teach her and to speak to her in the Alpha language. Those children came from mixed families but were not in fact bilingual, their knowledge of language Alpha was basic and mostly passive (they understood it but were reluctant to speak).

The other problem was the lack of willingness of the parents. Some seemed genuinely interested, but our communication ended with a questionnaire they filled in. Some families said openly that they do not have time for the cooperation or that the school is too tiring for the child as it is, so they do not want to give them extra work. Some families did not want the researcher to come into their home, saying that it is their private territory. Since all the conversations were to be recorded in children's homes where they feel the most comfortable, this parental attitude made it impossible to work with their children. Some of the researcher's e-mails, after a first optimistic phone call or a chance initial meeting, were left unanswered.

## 2. Data collection

The first stage of data collection was a questionnaire filled in by the parents. The document contains 50 questions about the child's personal data, language strategies used by the parents, their attitude to bilingualism, the child's contact with both languages, sources of input and the child's way of using both languages. The information was

always collected before the meeting with the child. Before the recordings, the parents also signed a consent form, based on the one used by E. Lanza (2004).

The data are audio recordings collected with a digital recorder. We want to investigate the way children talk about a specific subject, therefore we decided that carrying out the recordings in a spontaneous conversation with the children would be too risky. Spending long hours recording every child, we would have to wait for a long time for the structures concerning space and movement to naturally occur in the conversation. We had to turn to a different type of source of bilingual data – elicited information in an experimental setting. Given that we were working with young children, who in most cases did not know the researcher at all, the term *experimental setting* should be understood in a relatively loose way. One of the main concerns was that the children feel as much at ease as possible. That is why all the recordings were carried out in the children's homes, at their convenience, with their parents present in the neighbouring room.

The recorded conversation is based on a description of two sets of pictures. There are 26 pictures to talk about the spatial relations and 27 to talk about the motion events. Following M. Hickmann (2012: 28), we distinguish *voluntary motion*, in which the person or the animal moves in a certain way and on a certain trajectory, and *caused motion*, in which a person or an animal makes an object move on a trajectory and in a certain way. The total number of situations to be described by the child is 67. The child answers to two types of questions: “where is he/she?” and “what is he/she doing”? The work with the two sets of pictures is interspersed with some playtime of the child's choice (colouring, board games, jigsaw puzzles, etc.)

The researcher meets the children twice, once to talk about the pictures in French or English and the second time to talk about them in Polish. Given that the latter is the stronger language, the researcher first wants the child to believe that she knows no Polish. The parents present the researcher to the child as a French or English speaker and she does not say anything in Polish during the first meeting and recording. This of course can work only if the child has not met the researcher before, which is not always the case. For children who do know her, the researcher usually tells them that she has a special switch, which makes her speak only the Alpha language (a magic click of the fingers or pulling her ear). So far all of the children were willing to play the game and have cooperated by speaking only in the required language.

The children do not know what type of behaviour we are looking for, they are just told to look at the pictures and answer the questions or describe a part of the picture. Sometimes instead of answering the question in a descriptive way, they just use gestures and show something in the picture (“where is he”? “here”). In such a case they are usually told to describe it to a stuffed toy whose back is turned and who cannot see the picture, or the researcher explains that someone listening to the recording will have to draw the same picture and that is why they have to say where exactly the person is.

The duration of the recordings varies from 30 minutes to 1 hour and 30 minutes, depending on the child's willingness to cooperate, their fluency in the given language and on the length of the playtime between the two sets of pictures.

The problem we encountered concerns too long gaps between the recordings in the Alpha language and in language A. Both cannot be done on the same day – we do not

want to overstrain the children and, on the other hand, we want them to forget the pictures before they see them again. The gap between the two meetings should be at least a day or two, up to two weeks. Unfortunately keeping to this schedule has not always been possible, because of different factors, such as: the child's sickness, family travel or child's (or parents') tight every day timetable. Therefore, a few of the recordings had to be repeated.

### 3. Data transcription

The transcriptions are made in the CLAN program, a useful and rich tool for transcribing and analysing bilingual data, requiring formatting in CHAT Transcription Format (B. MacWhinney 2014).

#### 3.1. Utterances

One of the basic units of speech for transcription in CHAT is the utterance (the smaller ones being the morpheme and the word). As E. Lanza (2004: 123-124) points out, the term *utterance* can be understood as “a stretch of talk by one person, before and after which there is a silence” and it “should have a single intonation contour and single breath group”. M.T. Turell and M.G. Moyer (2010: 201) notice however that „there is no fixed way to define where an utterance begins or ends. It is open to the researcher to suggest the criteria to be used in order to define such a unit with respect to their data”, the criteria could be syntactic or intonational.

Without questioning the definitions above, we would like to point out however that a semantic criterion also seems to be operational in the matter of splitting the transcribed speech into utterances. When a fragment of speech contained a longer or shorter pause, but it expressed the same thought, it seemed more suitable to transcribe it as one single utterance, with the pause marked within, with (.), (..) or (...) symbols, depending on the pause's length (see example 1).

##### Example 1

- \*JAD: it's a sledge.  
 \*ADM: and we've got sledge (.) but you know this, with a staring  
 [: steering] wheel.  
 \*JAD: &uuu it's a funny sledge, you can steer it.  
 \*JAD: okay and now tell me what is this boy doing (.) you see, the boy  
 with this blue hat?  
 \*ADM: going down.

On the other hand, since a question mark always ends an utterance, each question needed to be transcribed as a separate utterance, even if they both related to the same subject (see example 2).

##### Example 2

- \*JAD: and where is Lola here?  
 \*ALX: in her house.  
 \*JAD: &mmhm.  
 \*ALX: which wooks [: looks] alike a twee [: tree] (..) trunk.

\*JAD: &oh it does, you are right it does look like [/] like a tree trunk.

\*JAD: but where is she exactly in the house?

\*JAD: <why can> [/] we can see her, why can we see her?

\*ALX: she's (..) we can see her from a window.

### 3.2. File constituents

Each CLAN transcription file written in the CHAT format has three main constituents: the file headers, the main tier and the dependent tier.

#### 3.2.1. File headers

##### A) Initial headers

CHAT has 7 initial headers of which 6 are obligatory in each CHAT file. They are the first visible elements of the file, the headers before them being hidden.

The headers: `@Begin`; `@Languages` (using languages codes from the international ISO 639-3 standard); `@Participants` (in that line the three-letter speaker ID for all the actors of the recording are introduced); `@ID` (for each participant) and `@Media` (used to identify the media linked to the transcript) are placed at the beginning of the file (see example 3). The `@End` header finishes the file and is placed at the end. (B. MacWhinney 2014: 25-30)

##### Example 3

`@Begin`

`@Languages: fr, pol`

`@Participants: MIE Mieszko Target_Child, JAD Jadwiga Investigator, LOA  
Lohann Father, EWA Ewa Mother`

`@ID: pol, fr|change_corpus_later|MIE|4;8.|male||Target_Child|preschool||`

`@ID: pol, fr|change_corpus_later|JAD||female||Investigator||`

`@ID: fr, pol|change_corpus_later|LOA||male||Father||`

`@ID: pol|change_corpus_later|EWA||female||Mother||`

`@Media: Mieszko_fr, audio, unlinked`

The `@Options` header was not used in our transcripts since suspending any checking rules of the program was not needed.

#### 3.2.2. Constant headers

The constant headers follow the obligatory initial headers. Their choice depends on the needs of the transcriber, they describe general facts about the file and are all optional (B. MacWhinney 2014: 30-32). The constant headers chosen for the transcripts are: `@Location` ; `@Time duration` ; `@Transcription` (see example 4).

##### Example 4

`@Location: Kraków, Poland`

`@Time Duration: 17:50-18:40`

`@Transcription: partial`

### 3.2.3. Changeable headers

The changeable headers, which can occur at the beginning or in the main body of the file, contain information that is prone to change during the transcription. They appear in the body of the file when that change takes place. Their list is long and includes headers like @Activities, @Comment, @New Episode, @New Language, @Page, etc. (B. MacWhinney 2014: 32-25)

Example 5

@Date: 05-MAR-2014

@Situation: investigator talks to the target child in the living room, the rest of the family is present: his mother, father and baby brother. The boy does not know the investigator who speaks to him only in French from the beginning so he does not know that she knows Polish as well. This part of the transcript is the first recording that day, it concerns movement and lasts up to 18 minutes and 60 seconds.

### 3.2.4. The main tier

The main tier is composed of utterances by the actors of the recording. Each utterance finishes with a full stop, a question mark or an exclamation mark. It is possible to code directly in the main tier the non-standard forms and pronunciation, the paralinguistic material as well as code-switches and errors made by the speaker.

#### A) Paralinguistic and non-linguistic material

Children's speech is filled with various paralinguistic behaviour. The speakers, young and adult alike, trail-off and do not finish their sentences, they make sounds such as: coughs, laughs, sneezes, gasps or add communicators such as “mmhm” or “uhmm”. Sometimes their words are accompanied or replaced by gestures. The CHAT format gives a possibility to transcribe this material in the main line of the transcript in a way so the CLAN program can treat those fragments as paralinguistic material and not as separate words. The following examples show how to code some of the paralinguistic behaviour:

- Trailing-off and communicators (example 6).

Example 6

JAD – Jadwiga, Researcher ; ALX – Alex, Target Child (6;1)

\*JAD: and this man here?

\*JAD: he's also holding a bag, he bought something +...

\*ALX: and he's coming out of a baby-shop.

\*JAD: &mmhm.

- Events like laughs or gasps used instead of words, in this case, the paralinguistic behaviour must represent a single utterance (examples 7 and 8).

Example 7

JAD – Jadwiga, Researcher ; ALX – Alex, Target Child (6;1)

\*ALX: I hate that because mummy has always too much of that stuff.

\*JAD: &=laughs.

\*JAD: and you don't like that?

\*ALX: no, she spends always most of the day before she has the babies is buying it.

#### Example 8

JAD – Jadwiga, Researcher ; NIC – Nicole, Target Child (7;4)

\*JAD: and the children here (.) where are they sitting?

\*NIC: &hmmm under the table.

\*NIC: &=gasps.

\*NIC: no, on the table.

- Paralinguistic events used while speaking – since the CLAN program automatically applies the code to the word directly preceding it, if a series of words is said in a specific way (for example, whispered), the whole passage has to be put in angle brackets (see example 9).

#### Example 9

JAD – Jadwiga, Researcher ; ALX – Alex, Target Child (6;1)

\*ALX: <how do they call this bear, because I forgot>[=! whispering].

\*JAD: Winnie [=! whispering].

\*ALX: and they're taking Winnie out of a hole.

- Non-linguistic behaviour – like in the case of paralinguistic events, if used instead of words, it forms a separate utterance. Example 10 shows a way of coding a head movement meaning “no”.

#### Example 10

JAD – Jadwiga, Researcher ; ADM – Adam, Target Child (4;2)

\*JAD: okay (.) and now tell me where is (.) her (.) you know what her name is, you remember?

\*ADM: &=head:no.

\*JAD: no?

\*JAD: Lotta.

#### B) Non-standard forms

The non-standard forms in children's speech might be of morphological or phonetical nature. They might also be a sign of interference between the two languages spoken by the child. Some forms however are a result of a widely used pronunciation which does not correspond with the written form of the word. The CHAT manual gives a list of spellings for this spoken versions of often used words like *yeah* for *yes*, *gimme* instead of *give me*, etc. This spelling has to be completed by the standard form of the written word put in brackets and following the non-standard version (see example 11).

#### Example 11

JAD – Jadwiga, Researcher ; ALX – Alex, Target Child (6;1)

\*JAD: and here?

\*JAD: I know you don't like that story ,, do you?

\*ALX: yeah [: yes].

The same procedure is used when the speaker mispronounces a word or uses an incorrect form. It is possible to write the incorrect form in standard spelling, if feasible (example 12) or to use the phonological coding (example 13).

#### Example 12

JAD – Jadwiga, Researcher ; ALX – Alex, Target Child (6;1)

\*JAD: and the fox, what is the fox doing?

\*ALX: the fox <was walking> [//] is walking over a stone and his sleeping bag fallled [: fell] into a puggle [: puddle].

#### Example 13

JAD – Jadwiga, Researcher ; MIE – Mieszko, Target Child (4;8)

\*JAD: les enfants ici, tu vois, ils boivent, ils mangent et où ils sont?

\*MIE: à la maison ils māz [: mangent].

If the non-standard form can be related to the particular child, it is possible to code it as a child-invented form, marked by the @c which follows it directly and an explanation is added in the comment tier (see example 14).

#### Example 14

ALX – Alex, Target Child (6;1)

\*ALX: no, he buyght@c I think some baby stuff.

%com: the form *buyght*, phonetically/bajt/, is a mixture of *buy* and *bought*

Another type of non-standard form of a word is the unfinished word. In order to let the CLAN program count the unfinished words altogether with the finished ones, the unpronounced ending is put in brackets (see example 15).

#### Example 15

ALX – Alex, Target Child (6;1)

\*ALX: becau(se).

\*ALX: I hate that because mummy has always too much of that stuff.

### C) Discourse repetitions and retracing

The speakers often repeat their words, or whole sentences without changing them. If the repetition concerns a single word, the symbol [/] should follow it. If it concerns the whole sentence, the latter should be put in angle brackets (B. MacWhinney 2014: 73), see example 16.

#### Example 16

JAD – Jadwiga, Researcher ; ALX – Alex, Target Child (6;1)

\*JAD: so what is he doing?

\*ALX: he's picking up <a box> [//] a box, he's picking up a box or a kind of (.)a heavy &um +...

\*JAD: it's a box, I think we can say it's a box.

When the speaker repeats the words to change them or to correct the sentence, but sticks to the same idea, the retracing symbol [//] is used (B. MacWhinney 2014: 74), see example 17.

#### Example 17

JAD – Jadwiga, Researcher ; ALX – Alex, Target Child (6;1)

\*JAD: and the fox, what is the fox doing?

\*ALX: the fox <was walking> [//] is walking over a stone and his sleeping bag falled [: fell] into a puggle [: puddle].

#### D) Code-switching

Working with bilingual children's recorded material often gives the researcher examples of a code-switch. For each file, the main language of the recording is given as the first one in the @Languages header. The second one is the other language spoken by the child, but it is not the language of the conversation. In the collected material, the code-switching seems to appear mostly from the Alpha language to the language A (the stronger of the two), when the child does not remember, or know the right word. The switch is marked in the transcription with a pre-code. In example 18, the conversation is in French and the researcher is not using Polish. Not remembering how to say “street” in French, the Target Child switches to Polish, which is marked by [- pol] preceding the utterance.

#### Example 18

JAD – Jadwiga, Researcher ; HEL – Helena, Target Child (4;8)

\*JAD: tu vois ici?

\*HEL: grenouille!

\*JAD: oui, où elle est la grenouille?

\*JAD: qu' est ce qu' elle fait, où elle est?

\*JAD: ça c' est quoi?

\*HEL: [- pol] ulicia [: ulica].

\*JAD: oui, c' est la rue.

\*HEL: la rue.

#### E) Unidentifiable material

Because of the surrounding noises, overlapping speech or sometimes because of the child's volume of speech, some of the material turns out to be impossible to identify. Those fragments are marked in the transcription by the “xxx” code (see example 19). Although, in some cases it is possible to at least take a guess of what the child is saying, this “best guess” can be given in the comment tier or coded with a question mark at the end of the utterance (B. MacWhinney 2014: 71–72), see examples 20a and 20b.

#### Example 19

JAD – Jadwiga, Researcher ; HEL – Helena, Target Child (4;8)

\*JAD: &ah oui, t' as fini ici (.) et tu peux me dire où elles sont?

\*HEL: xxx xxx [=! whispers].

\*JAD: Helena.

\*HEL: quoi?

- \*JAD: tu peux me dire où elles sont?  
 \*HEL: je sais pas (.) &ah sur la neige.

#### Example 20a

JAD – Jadwiga, Researcher ; ADM – Adam, Target Child (4;2)

- \*JAD: and this girl here, what is she doing?  
 \*ADM: xxx xxx on the road.  
 %com: probably "riding bicycle"

#### Example 20b

JAD – Jadwiga, Researcher ; ALX – Alex, Target Child (6;1)

- \*JAD: i tutaj mamy jeszcze, gdzie jest dziewczynka?  
 \*ALX: na ściesce [: ścieżce] po drodze do domu [?] a mama jest z tyłu.

### 3.2.5. Dependent tiers

The dependent tiers contain information about the utterance above such as: codes, comments, events, and descriptions of interest to the researcher. They make the transcript more readable as keeping all the extra information coded in the main line would make the file “heavier” and more difficult to read. They can be used on the transcribing level or on the coding level (B. MacWhinney 2014: 78–79). There is a long list of standard dependent tiers from which we are advised to choose when creating a transcript file in CLAN, it is also possible to create one’s own codes. The examples show as follows:

- Paralinguistic Tier (paralinguistic material can be coded in both main and dependent tiers, depending on the transcriber’s decision and on the clarity of the main tier).

#### Example 21

JAD – Jadwiga, Researcher ; MIE – Mieszko, Target Child (4;8)

- \*MIE: &o regar [: regarde].  
 \*JAD: &mmhm.  
 \*JAD: il est sur l' arbre, il les regarde, maintenant il vient les aider.  
 \*MIE: 0.  
 %par: laughs  
 \*JAD: ça c' est rigolo ,, oui?

- Action Tier

#### Example 22

JAD – Jadwiga, Researcher ; MIE – Mieszko, Target Child (4;8)

- \*JAD: il court, et avec le caillou, qu' est ce qu' il fait, avec la pierre?  
 \*MIE: comme ça.  
 %act: Mieszko jumps  
 \*JAD: tu sais comment [/] comment on le dit en français?  
 \*MIE: non.  
 \*JAD: il saute.  
 \*MIE: il saute?

- Addressee Tier, used when the child talks to a different person than the researcher.

## Example 23

JAD – Jadwiga, Researcher ; ALX – Alex, Target Child (6;1) ; ADM – Adam, Brother (4;2)

\*JAD: so now I'm gonna [: going to] work with &uh Alex.

\*ALX: <could you> [/] could you come out of the room?

%add: ADM

\*ADM: [- pol] mogę się tylko psytulić [: przytulić]?

- Comment Tier, used for general purposes and can contain different types of researcher's comments (i.e. a child's behaviour, an attitude, a transcriber's interpretation of child's words, information about the accent).

## Example 24

JAD – Jadwiga, Researcher ; ADM – Adam, Brother (4;2)

\*JAD: where is this boy?

\*ADM: outside!

%com: Adam is getting annoyed

## Example 25

DAW – Dawid, Target Child (6;4)

\*DAW: je faité [: fait] ça.

%com: „fait” or „fait tout”

The problems and doubts arising during the transcription phase were mostly related to the „baby pronunciation” that some of the children still had. Those aspects of children's speech were treated as other examples of non-standard pronunciation. The signs of children's bilingualism (i.e. the accent) were commented on in the dependent tier line. The comment line was also used to explain other non-standard forms appearing in the recordings, were they errors made in one of the languages (because of insufficient level yet) or being signs of interference from one language to the other. A problem which appeared only when transcribing recordings in French were the homophone forms of the verbs, when the infinitive, the past participle (*participe passé*) and the *imparfait* past tense forms had the same or very similar pronunciation. It was not always obvious which one of those forms the child was using. The alternative spelling and comments were added in the comment tier line. In example 26 the pronunciation was rendered and the correct form (which should be used in this context, although pronounced differently) was given in brackets, and in example 27 the form in the brackets was chosen because in other cases of the phonological form *seve* it was always followed by the infinitive. We believe the latter to be the changed form of the modal verb *vouloir* (*il veut*)

## Example 26

MIE – Mieszko, Target Child (4;8)

\*MIE: &mmhm (.) il rule [: roule] à vélo.

%com: /rule/ (rouler/ roulé/ roulais) - here t

Example 27

MIE – Mieszko, Target Child (4;8)

\*MIE: non, &seve &yr c' est un garçon, mais &seve zâtre [: entrer] là.

%com: [zâtre] (entrer, entré, entráis)

## 5. Conclusion

Conducting any research involving bilingual children poses many challenges at the methodological level, both at the stage of data collection and data transcription.

### (1) Data collection challenges

The first difficulty we were faced with was establishing the criteria for the selection of candidates, which is linked to the understanding of the term bilingualism itself (in our research treated in the narrow sense of the term). The next step was finding the right candidates, which was not an easy task, given the fact that not all of the presumably bilingual children met the criteria. Also, not all of the families wanted to cooperate, and those who wanted did not always have the time for it.

Also, children's bilingualism, even within a single group of participants having met the established criteria, turned out to be very diverse. Depending on the type, source and load of input in the Alpha language, the level of proficiency in French or English was uneven, even among children of the same age. It meant that more flexibility was called for – to increase the child's motivation or to put the young speaker on the right track as for the required answers to the researcher's questions.

### (2) Data transcription challenges

The transcription itself, thanks to the CHAT format, is very operational. The format, paired with the CLAN program gives possibilities for different types of speech data processing – simple ones and those with big amounts of extra information coded. It also lets the researcher choose what information he wants to include, transcribe and code, according to the needs of the given research, which makes it a universal tool. It also allows to upload the data into the CHILDES Internet Database and make it useable for other researchers in child speech and bilingualism.

Encountered difficulties, especially those relevant to data collection, show the importance of the studies about children's bilingualism – some parents' knowledge about the phenomenon is full of prejudice and misinformation, they are sometimes afraid of affecting their child's speech development or just not thinking that bilingualism is worth fighting for. We dare to hope that the conducted research will contribute to filling the gap in existing studies in the field of Polish bilingualism.

## References

- Błasiak, M. (2011), *Dwujęzyczność i penglish. Zjawiska językowo-kulturowe polskiej emigracji w Wielkiej Brytanii*. Kraków.
- Davidson, C. (2010), *Transcription matters. Transcribing talk and interaction to facilitate conversation analysis of the taken-for-granted in young children's interactions*. In: *Journal of Early Childhood Research* 8, vol. 2, 115–131.

- De Houwer, A. (2009), *Bilingual First Language Acquisition*. Bristol.
- Hickmann, M. (2012), *Diversité des langues et acquisition du langage: espace et temporalité chez l'enfant*. In: *Langages* 188, 25–39.
- Hoffmann, C. (1991), *An Introduction to Bilingualism*. London.
- Košťalova, K. (2010), *Mowa dwujęzycznej Julki – studium przypadku* (Master's dissertation, Jagellonian University).
- Lanza, E. (2004), *Language Mixing in Infant Bilingualism. A Sociolinguistic Perspective*. Oxford.
- Laskowski, R. (2009), *Język w zagrożeniu. Przyswajanie języka polskiego w warunkach polsko-szwedzkiego bilingwizmu*. Kraków.
- MacWhinney, B. (2014), *The CHILDES Project. Tools for Analyzing Talk - Electronic Edition, Part 1: The CHAT Transcription Format*. Carnegie Mellon University. <http://childes.talkbank.org/manuals/chat.pdf>, website visited 02/03/2015.
- Mróz, M. (2011), *Bilingual Language Acquisition: Focus on Relative Clauses in Polish and English*. Warszawa.
- Paprocka-Piotrowska, U./ G. Komur-Thillooy (2012), *Jamais deux sans trois? Comment l'enfant s'approprié trois langues parallèlement (étude de cas)*. In: *Studii di lingvistica* 2, 123–141.
- Turell, M.T./ M.G. Moyer (2010), *Transcription*. In: L. Wei/ M.G. Moyer (eds.), *The Blackwell Guide to Research Methods in Bilingualism and Multilingualism*. Oxford, 192–213.