The Composition of the Bilingual Lexicon

Annabelle David and Li Wei
University of Newcastle-upon-Tyne

Introduction

Research on bilingual children's lexical development has focused on issues such as the size of the vocabulary in each language, the relationship of the two lexicons (separate or "fused"), and the development of translation equivalents. There has been little research on the nature of the lexicon itself, i.e. the composition of the developing bilingual lexicon. Cross-linguistic studies of monolingual children speaking different languages consistently claim that all children follow a similar developmental pattern in that different lexical categories develop at different times. For example, Benedict (1979), in a study of early production and comprehension of 8 English-speaking monolingual children, found that 61% of the first 50 words were nominals. 19% were action words (social games, locatives, verbs…) and 10% modifiers. Although Benedict's categorisation is debatable, the results show that nouns are clearly the first type of words to appear and the most common category in early lexical development. Bates, Marchman, Thal, Fenson, Dale, Reznick, Reilly and Hartung (1994) conducted one of the first large-scale study of lexical development based on the American *MacArthur Communicative Development Inventory* (CDI). Their results determined three "waves" in lexical acquisition. The first wave, to the 100-word stage, is characterised by an increase of common nouns in the child's vocabulary. The second wave sees a decrease of the percentage of nouns in the lexicon and an increase in the percentage of predicates (mainly verbs). This phase takes place when the child has between 100 and 400 words in his/her lexicon. Finally, the third wave, over 400 words, is a sharp increase of closed-class items, stable until now.

Caselli, Casadio and Bates (1999) studied monolingual Italian children based on the Italian adaptation of the CDI. They proposed a four-stage model of lexical development. Those four stages are *Routines and Word Games, Reference, Predication and Grammar*. From the onset of speech up to the acquisition of about 10 words in his/her productive vocabulary, the child is said to be in the *Routines and Word Games* stage. During that time, children acquire social words, i.e. items like *bye bye* or animal sounds. Very often it is quite impossible to classify those words into grammatical categories. They are referred to as social words as they allow the children to have a social interaction with their environment. The second stage (*Reference*) is when the child has between 50 and 200 words. That stage corresponds to the rapid acquisition of nouns. Those nouns, most of the time, refer to concrete objects: *table, cat, dress*, etc. The third stage sees the appearance of verbs and adjectives (predicates). It may start as early as the 100-word mark. Stages 2 and 3 actually overlap. Finally, during the last stage, children acquire "grammatical words" (e.g. pronouns, prepositions etc.), rare until then. This stage also correlates with the emergence of productive grammar. Eriksson (2001) reported the same 4 stages for Swedish monolingual children in a study based on the Swedish adaptation of the CDI. Other results for Spanish-speaking children (Jackson-Maldonado, Thal, Bates, Marchman & Gutierrez-Clellen 1993) and Hebrew (Maital, Dromi, Sagi & Bornstein 2000) have confirmed those stages.

Cross-linguistic studies of such kind have important implications for research on the developing bilingual lexicon. Do bilingual children, acquiring two languages simultaneously from birth, have the same/similar developmental patterns as their monolingual counterparts. This paper addresses this question with a study of the composition of lexicon of 13 French-English bilingual children between the age of 1;0 and 3;0. We study the development of the different grammatical categories based on previous cross-linguistic studies of monolingual children and compare the development in each language. Before we look at the methodology and data, a brief review of the existing studies of lexical development of French-speaking children is in order.

1 French-speaking children's lexical development (and other Romance languages)

Bassano, Maillochon and Eme (1998) reported no developmental progression of the lexical diversity of a single French child. All the categories were present almost from the start. As expected, they reported that nouns largely dominate until 1;8. After that age, the proportion of nouns in the lexicon decreased to make space for predicates and "grammatical words" or closed-class items. The first category had a slow and steady progression, while the second one seemed to have a more dramatical increase after the age of 2;0.

A French version of the CDI has recently been developed. The first study using the French CDI is still under way. However, preliminary results tend to generally concur with those reported for other languages. Kern (2001) found that the noun category increases until the 100-word stage and then remained quite stable until the lexicon reached 500 words. Predicates increase regularly as a function of the size of the lexicon. It was also found that children with a vocabulary of less than 50 words tended to produce as many closed-class items as predicates.

While all developmental studies on lexical acquisition seem to agree on the overall distribution of grammatical categories, some minor differences can be found between Romance languages, particularly French, and English, that will nevertheless impact on the research design and subsequent findings. A very interesting finding from Bassano et al. (1998) is the fact that the closed-class category is proportionally always much larger in French than the same category in English-speaking children. According to Bates et al. (1994), the closed-class category never represents more than 15% of the whole lexicon until at least 30 months in English. Whereas, Bassano et al. (1998) found that at 1;2 (when she had less than 50 words in her lexicon), Pauline, the child they studied longitudinally, had 20% of closed-class items in her total lexicon. In a comparative cross-sectional corpus, Bassano et al (1998) found that at 1;8 and 2;6, between 8% and 45% of the French-speaking children's lexicon were closed-class items (above 20% on average).

Another study of a romance language seems to confirm Bassano's (1998) results. Caselli et al. (1999) compared Italian to English lexical data both based on the Communicative Inventory. They also found that the proportion of the closed-class category in the lexicon was always slightly higher for Italian children.

A possible reason for the apparent difference would be that French, and other Romance languages like Italian, do contain a lot more closed-class items. For example, for a single word like the possessive *my* in English, there are at least three possible translations in French: *mon*, *ma, mes*, which are respectively, the masculine, feminine and plural forms. Italian has *mio*, *mia, miei* and *mie*. Spanish has two forms *mi* and *mis*.

The predicates categories contain slight differences in their developments and quantities too. English has a basic SVO word order. Although French usually has an SVO structure, it sometimes changes to SOV order when the object is pronominalised (E.g.: *Je l'ai mangé*). French also has an extremely rich system of verb morphology and verb agreement plays a crucial role in conveying basic sentence relations compared to English. So theoretically speaking, French could provide a basis for earlier verb acquisition or larger proportion in lexicon. Bassano et al. (1998) found a slight difference between English and French. They reported a slightly larger number of predicates initially. Predicates in the French-speaking children's lexicon represented on average 18% when around the 50-word stage. They only represented 8% for the American children according to Caselli et al. (1999). Similarly, Boysson-Bardies (1999) found that French children with a lexicon inferior to 50 words had 21% of verbs, while American children had only 9%.

However, a study on another Romance language with a rich verbal morphology did not show the same results at all. Italian is a so-called pro-drop language where the verb may be produced without an explicit referential or pronominal subject constituent. So that the subject can be dropped and the first word becomes the verb. In spite of its important role, Caselli et al. (1999) found no evidence to prove that verbs would appear earlier in Italian children. The data showed that up until the 300-word stage, Italian-speaking children had proportionally less predicates in their lexicon, even if the difference was minimal at the later stages.

Hebrew is a language where the word order is relatively free and where verb initial constructions are commonly found in everyday speech. However, Maital et al. (2000) did not find any conclusive evidence that predicates (or verbs) appeared earlier or were more prominent in Hebrew compared with English (respectively 4% and 8% at the below-50-word stage). No similar research has been done with bilingual children. It would be interesting to know whether those cross-linguistic differences are present in the bilingual lexicon: e.g. whether French-English bilingual children produce more verbs and closed-class items in French.

2 Aims and objectives

Our first task will be to determine whether the overall lexical development in bilingual children follows the same patterns as that of monolingual children, whatever language they may be acquiring. In other words, we will try and demonstrate whether bilingual children's vocabulary is composed of the same types of words (verbs, nouns, etc.) at the same given age or rather for a same lexicon size. Do bilingual children go through the same stages? Perhaps more importantly, we will also try to examine if the two languages evolve in a parallel way or if they each have a different developmental pattern, i.e. whether the two lexicons follow a similar development or not in their distribution (and quantity). For example, do nouns develop in both languages at the same time or does one language contain most of the nouns and the other one the other categories?

The language dominance factor might play a role in the way the lexicon is built in bilingual children. If a child is clearly dominant in one language (lexically speaking), that dominant language might contain elements from all categories (nouns, verbs, prepositions, etc.). But the other language, not being as developed, might only contain some elements; maybe those that appear first in language acquisition, namely, social words and nouns.

If a child then has all the necessary elements in one language to build a sentence, but only some in the other, then that would lead a bilingual child to a need for code-switching. In order to try and answer all these questions we will study each language separately and then compare them together.

It has been recognised that normative guidelines based on monolingual children give inaccurate predictions for bilingual populations. Comparisons of number of words should therefore be avoided (Pearson 1998). That is why we will try and avoid this kind of comparison. The aim here is not a quantitative study of the composition of the bilingual lexicon but rather a qualitative one, although we shall use quantitative methods for some of the analysis.

Our aims, therefore, will be to:
- Test to what extent the developmental lexical patterns observed for monolingual children are true for bilingual speakers.
- Examine whether the stages follow the same distribution and evolution in both languages.
- Check for cross-linguistic differences between French and English.

3 Analysis
3.1 Participants

The data is taken from a longitudinal study of 13 bilingual French/English children aged between 1;0 and 3;0. Data was collected in two countries, the U.K. and France. As it is often the case in developmental research, all the families were from a middle-class background. The families were asked to supply data about their child's health history, and their own education, occupation and other questions. The families were followed for, on average, two years from when the child started to talk. One condition for participating in the study was that one of the parent was a native French speaker and the second one a native English speaker. All of the French speakers were native monolingual French except for two who grew up bilingual. However, they were not using any of the second language with their child (Italian and German). All of the English speakers were British English native speakers but two who were Australian and New Zealander. All the parents had at least basic knowledge in the other language. All of the children were being brought up using the one parent - one language strategy. All of the parents acknowledged using that strategy more or less strictly. An effort was made to obtain a

similar number of female and male subjects as well as subjects living in France or in the U.K. Table 1 shows the distribution of the participants (all names used are pseudonyms).

Table 1. Children's linguistic characteristics

	Gender	Country of residence	French-speaking parent	English-speaking parent	Family language	Family ranking
THO	Male	France	Father	Mother	French	Only child
FLO	Female	England	Mother	Father	French	Oldest*
ANT	Female	England	Mother	Father	English	Only child
EMM	Female	France	Father	Mother	French	Youngest (out of 2)
DUN	Male	France	Mother	Father	English	Youngest (2)
PEN	Female	France	Mother	Father	English	Youngest (2)
ELI	Female	England	Mother	Father	English	Only child
JAC	Male	France	Father	Mother	English/ French	Youngest (2)
ANN	Female	England	Father	Mother	English	Youngest (2)
OLR	Male	France	Father	Mother	French	Oldest*
LIA	Male	France	Father	Mother	French	Only child
REB	Female	England	Mother	Father	English	Youngest (3)
OLB	Male	England	Mother	Father	English	Youngest (2)

*Younger sister born during study.

A very important point to be made here is that families were recruited through personal networks of the researchers (first author of this paper) who met with the families. They were all explained during the first visit what the study involved and what was expected of them. Those visits also allowed the researcher to observe first-hand the children and the families' language behaviour in situations, especially the use of the one person - one language rule. The visit often lasted for a whole day if not more. Thus, such daily activities as meals, play, bath, allowed a better understanding of the family's normal language choices and patterns. Regular contact was then kept throughout the study. On average, once a week, one parent would update the researcher on the child's progresses and the researcher would ask for any information and clarifications necessary for the study.

3.2 Materials

Most of the data for this study is based on the MacArthur Communicative Developmental Inventory (CDI). The original CDI is the American English version developed by Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick & Reilly (1993). The British English toddler version (ECDI) was adapted by Klee & Harrison (2001) for children between 16 and 30 months old. It contains 672 words organised into 22 semantic categories. The French toddler version (FCDI) called *Inventaire Francais du Developpement Communicatif chez le nourrisson: mots et phrases*, was adapted by Kern (submitted). It contains 689 words organised in the same 22 semantic categories. Both these adaptations are based on the original American version.

The CDI is one of the best parent report tool to analyse children's vocabulary. Its reliability has been demonstrated and validated (Dale, Bates, Reznick & Morisset 1989). The adaptations used for the present study are also undergoing studies to prove their reliability (Robertson 1999, Marr 1999, Harrison 2001, Kern 2001, Kern submitted). Another advantage of the CDI is that it contains naturalistic productions to a certain extent. Onomatopoeias like *uh oh* or animal sounds are included. Bassano et al. (1998) underline the importance of these very frequent words in early lexical development.

However, some problems remain. A few words in the French version are literal translations of the American rather than adaptations as it is suggested they should be. The French vocabulary list contains unusual items like the word *étendage*. Our guess is that it is a literal translation from the American *dryer*. However, we doubt this word is ever used by French children if at all by anybody. Many parents have reported this word as "inappropriate". The problem might lie in the fact that French speakers adapted the CDI and only one English speaker (it is not mentioned if s/he was American) advised in case of a conflicting opinion. Kern (submitted) states that they "tried to stay as close as possible to the American version, for purposes of crosslinguistic comparisons". It was also reported by many of the parents taking part in this bilingual study that some words in the British adaptation sounded American English and would never be used by them.

In addition to the problems associated with the CDIs, a further problem arises when wanting to use them with bilingual speakers. They are originally designed to be used with monolinguals. A few studies have used them with bilingual children. Among those, Pearson, Fernandez & Oller (1995) used the original American CDI together with its Spanish adaptation on bilingual children. One of the issues that arises when asking parents of bilingual children to fill out these forms, is that they might be more likely than monolingual parents to inflate their child's vocabulary (Pearson et al. 1995). Indeed, very often they might attribute a word to one language that was actually produced in the other language. However, it was judged that this problem might mostly arise when trying to judge comprehension. As far as production is concerned, it is easier to be accurate as the parent can refer to a more concrete event. This can also be avoided by keeping a diary of the child's new words (this solution was adopted in this study for up to 50 words). Another issue when dealing with bilingual children is the problem of undetermined words. It has been acknowledged that bilingual children produce words that could belong to either language, especially earlier in their lexical development (Deuchar & Quay 2000). As far as the CDIs are concerned, if a child says something like *mama*, it is hard to judge from which language the word comes from. One could argue that this problem could also happen in the case of monolingual children. Children use one same word or sound that can mean many different things in the adult's lexicon.

3.3   Procedure

Each parent was asked to fill in the CDI monthly. The parents were sent (or given) each month (on the day of the child's birthday) a copy of the FCDI and ECDI with a return envelope. They were asked to fill them in as they arrived. The next questionnaire was not sent until the previous one had been returned in order not to influence the next replies.

Each parent would fill in the CDI that corresponds to the language he/she is speaking to the child. However, we noticed that there could be a difference between the one filled in by the mother who stays at home all day taking care of the child and the father's who works all day, for example. So it was decided that two checklists would be obtained whenever it was possible for each language. However, the process was not repeated every month. The data for this study combines both sets (native and non-native parents' lists).

The British English CDI was filled out also by the two parents who were not native British English speakers. They were both given the British English version and another version (respectively New Zealand and Australian) to compare. Both estimated that their spoken English was closer to British English and preferred its content. To my knowledge, there is no Swiss French adaptation to this day.

The data collection process started at 16 months and stopped at 30 months. Not all of the families filled in all of the questionnaires regularly, so the number of checklists is not the same for each child. The parents were warned of the typical mistakes that could be made by parents of bilingual children. They were explicitly explained that if they tick the word *cat* in English, for example, it means that the child actually produced the word *cat* and not *chat* in French. The total of CDIs filled out and returned is 364.

Table 2. Age (in months) at which at least one CDI in each language was returned.

THO	17	20	21	**22**	23	**24**	25	26	27	28					
FLO	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
ANT	16	17	18	19	20	21	**22**	**23**	24	25	28				
EMM	16	18	20	21	23										
DUN	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
PEN	17	18	21	22	**23**	26									
ELI	16	17	18	19	20	**21**	22	23	24	25					
JAC	16	17	18	19	20	21	22	23	24	26	27	29	30		
ANN	16	17	18	19	20	21	22	23	24	25	26	27	28	**29**	30
OLR	16	17	18	19	20	21	22	23	24	25	27	28	29	30	
LIA	16	20	23	27	30										
REB	16	17	19	**20**	21	23	24	26	**29**	30					
OLB	16	17	18	**19**	**20**	21	23	24	25	27	28	29	30		

Numbers in bold mean that only one out of the two CDIs was completed.

3.4 Data reduction

Both CDIs are organised into 22 semantic categories. They also have several grammatical questions. However, we will not consider the grammatical questions in this paper. In order to determine how the lexicon is organised we needed to classify the reported words into grammatical categories. Since the data is taken from the CDI, there is no contextual information. By classifying the words under grammatical categories, we do not want to say that children recognise grammatical categories like noun, verb or adverb, as there are some examples where they deviate from the adult pattern even for a same word. However, children words have largely been classified according to adults' part-of-speech (English or French) in the literature. The categories are seen as variables: due to the input received, children treat nouns differently than verbs and so on. It does not imply the child's ability to realise the emergence of such categories either. It only allows the researcher to point out cross-linguistic differences in the vocabulary composition.

The categories used are the same as those used by Bates et al. (1994) from data obtained thanks to the CDI. Those categories were also used to compare American and Italian data (Caselli et al. 1999). Similar categories were used by Dromi (1996): *object words, action/state words, modifiers, social words*. The only difference is that she created an *indeterminant* category to put in anything that did not fit into any of the other categories. The categories are as follows:

Common nouns:
  Animals
  Toys
  Vehicles
  Food
  Clothing
  Body parts
  Furniture and rooms
  Outside things
  Household objects

Predicates:
  Action words
  Adjectives

Closed-class category:
  Helping verbs
  Connecting words
  Quantifiers and articles
  Prepositions and locations
  Question words
  Pronouns

Social words:
  Sound effects
 People
  Games and routines

Two categories of the CDI were left out *Places to go* and *Time words*; the reason being that the words contained in these categories can be adverbs as well as nouns both in French and in English. Overall, this classification can appear arbitrary and arguable, especially for social words or expressions (e.g. *vroom* or *give me five!*). Consequently, cross-linguistic differences that can be over-turned by one or two items should be taken very carefully. These categories almost correspond to the function word/content word classification. Content words would approximately be common nouns and predicates. Although this classification was originally used in English, it was also successfully used in Italian, Swedish, Hebrew and others. It is also suitable for French. The only problem that might arise is the auxiliary category. Modals in English (*can*, *could*…) are not equivalent to their translated forms in French like *pouvoir, vouloir* and *devoir*. These French verbs behave like normal lexical verbs, especially, in terms of the inflections they take. The only true auxiliaries are *être* and *avoir* in compound tenses. The FCDI does not have many real auxiliaries in its auxiliary category. However, it was decided that the category would be kept within the closed-class items. Very young children produce very few auxiliaries (so the choice will not change dramatically the results). Because the categories do not aspire to be real grammatical categories recognised by the children, auxiliary category can be left in the closed-class part. Some verbs such as *aller*, *avoir* and *faire* are also in the *action words* category (therefore, also in the predicates). So the forms proposed as auxiliaries are thought to be taken as such: *fait*, *a fait*, etc. Moreover, these verbs do not carry the main semantic meaning. They are all to be attached to something else. Their role is largely if not wholly grammatical. We can consider auxiliaries as: verbs that help the main verb in expressing inflection (Jones 1996). Even if some are more modals than auxiliaries: e.g. *vouloir*. They do not carry the main semantic meaning. This category could actually be renamed inflection verbs. And as such they are to be classified with the closed-class items. They are simply non-action verbs (Bassano 2000). Finally, Kern (2001) made the same choice for French monolingual lexicon analyses, although they did not account for social words.

4   Results

Taking into account the high rate of variability among children at the same age, grammatical distribution was considered as a function of vocabulary size and not as a function of age. The entire sample was divided into seven categories according to the size of the produced vocabulary in that language (not the total of both languages): (1) less than 50 words, (2) 51-100 words, (3) 101-200 words, (4) 201-300 words, (5) 301-400 words, (6) 401-500 words and (7) more than 500 words. It is important to note at this stage that none of the children reached the last stage (more than 500 words) in their English lexicon. Therefore, the category is not represented on the following graphs.

Figure 1 (English lexicon) shows that the composition of early lexicon differs according to vocabulary size. Nouns constitute the most produced items, with the exception of children with a repertoire fewer than 50 words. We can notice several clear trends. The largest category at the below-50-stage is by far the social words category (63%). This category falls quickly to the benefit of nouns. Around the 100-word landmark, the common nouns category stabilises before slowly decreasing. And at the same time, the number of predicates and closed-class items start increasing. The first one rises quicker than the second one.

Figure 1. English lexicon
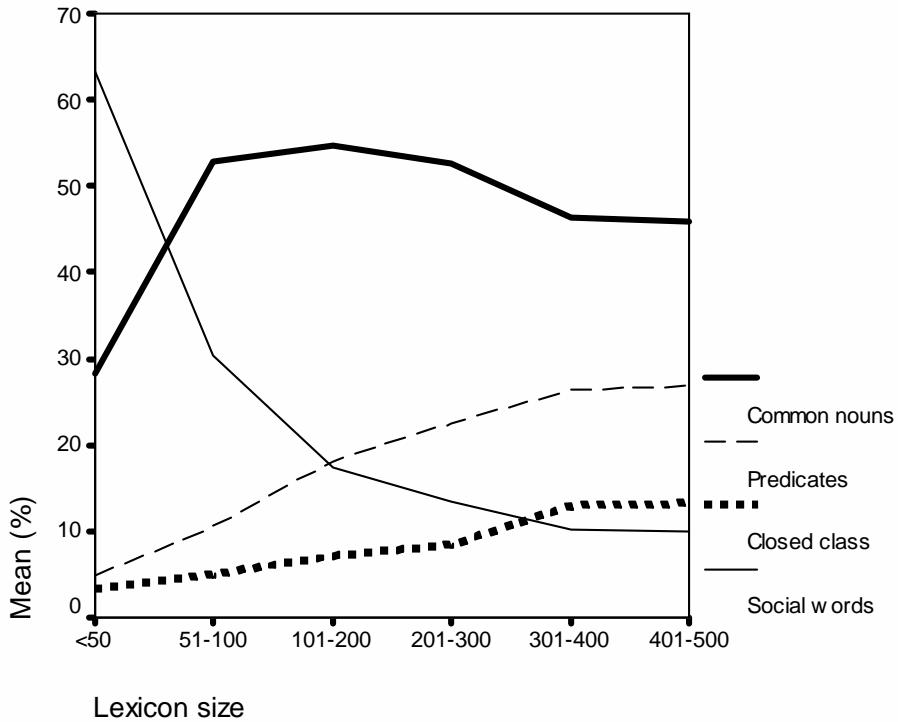
The graph representing the French lexicon (figure 2) of the bilingual children seem at first sight to be very similar to the English one. The social-word category represents most of the items produced for children with a lexicon below 50 words (58%). This category decreases rapidly. Nouns then increase quickly until about 100 words, stabilise and then start decreasing slowly. The decrease of nouns seems to trigger the increase of predicates and closed-class items. The results for both languages are consistent with monolingual data reported in previous studies.

Let us have a look at each category separately. The only noticeable difference between the two lines on figure 3 is the slightly higher number of predicates in the French lexicon for children with the smallest vocabulary (ECDI: 4.8%, FCDI: 7.3%). However, this trend fades away as the vocabulary increases and even inverses. From 50 words, the number of predicates rises much quicker in English than in French. At their highest point, respectively for a lexicon of 401-500 and above 500 words, predicates represent 26.8% in English and 23.4% in French. It is important to note that the ceilings of the predicates category are 24.7% for the ECDI and 24.2% for the FCDI.

A list was drawn for ten children of their first 10 words in order to investigate further this tendency. This list is based on detailed diaries kept by the parents and regular correspondence between the parents and the researcher. This list (in table 3) shows that out of the ten, only one child produced one verb in English. That verb (*cuddle*) was counted as a predicate. However, it could also be placed in the common nouns category. If the child produces that word in isolation, it is impossible to decide whether it is a noun or a verb. On the other hand, six children produced one or two verbs in French. No child produced a verb in each language. 1 child had an adjective among his first ten words (placed in the predicates category in the CDIs).
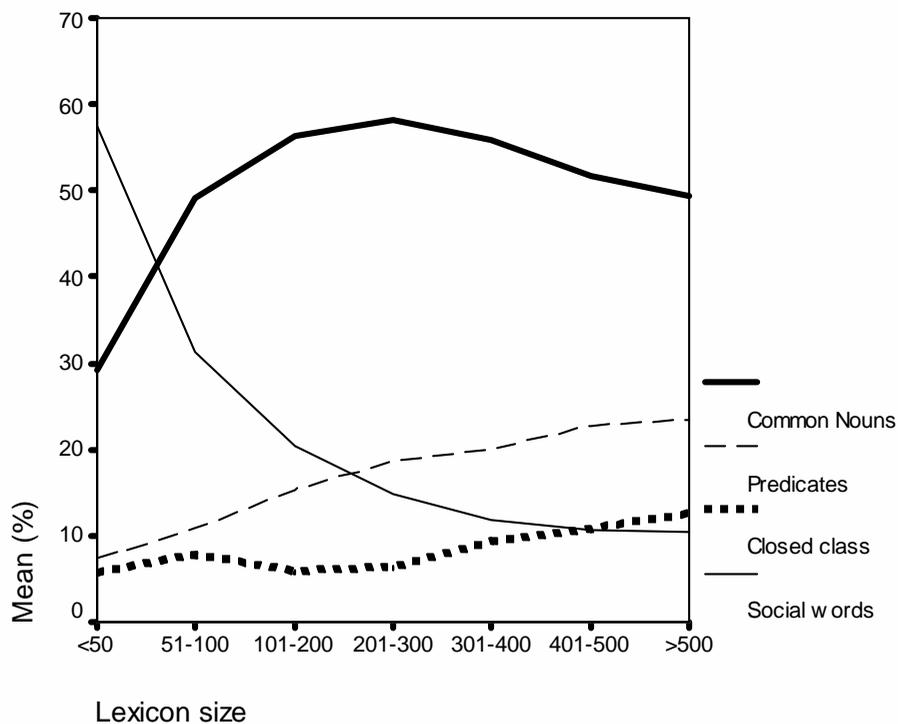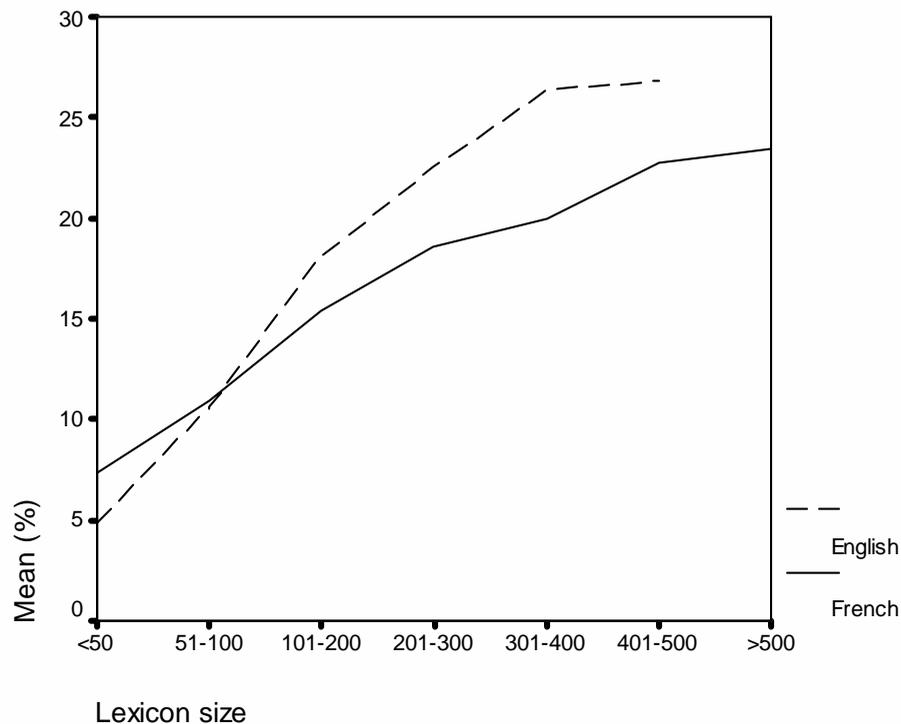
Figure 2. French lexicon

Legend:
- Common Nouns
- Predicates
- Closed class
- Social words

Figure 3. Predicates

Legend:
- English
- French

Table 3. First 10 words

FLO	JAC	ANN	EMM	REB
Dog	Mummy	Papa	Ball	Bye
Door	Papa	Mummy	Grandpa	Shush
Maman	Banana	Book	Mummy	Doudou
Daddy	**Boire**	Byebye	Byebye	Allo
No	Poisson	Hi	Shush	Bravo
Tiens	Bear	No	Thank you	Maman
Donne	There	Yes	**Cuddle**	**Entendre**
Shoe	Pain	Ball	More	Daddy
Crocodile	Aspirateur	Nez	Allo	Merci
Cat	Byebye	Bear	Papa	Non
ELI	**DUN**	**LIA**	**OLB**	**OLR**
Byebye	Maman	Apple	Allo	Au revoir
Daddy	Allo	Daddy	Encore	Daddy
Tiens	Hi	Hot	Cat	Maman
Maman	**Tomber**	**Attend**	Tracteur	Tickle
Merci	Banana	Oh dear	Coucou	Non/no
Encore	Dad	Maman	Salut	Bébé
Calin	Yaourt	Door	Shush	Coucou
Ball	**Boire**	Pardon	Caca	Pain
Hello	Canard	Gateau	Lait	Shoe
Baby	Byebye	Ball	Bras	Papi

Based on these two results, it seems that those children acquired predicates earlier in French than in English.

The closed-class category is about the only one that does not follow a parallel pattern in the two languages. This has been observed in Hebrew (Maital et al. 2000) and Italian (Caselli et al. 1999) also when compared with English. The proportion of French closed-class items decreases between 50 and 100 words before increasing again. This slight decrease between 50 and 100 words has also been found in monolingual French-speaking children (see Kern 2001). Here again, the French lexicon seems to contain more closed-class items for children with the smallest vocabulary. However, this trend inverses from about 100 words. The ceilings for this category are 15.1% and 13.4% respectively for the ECDI and FCDI. The bilingual children by 500 words reached on average the 12.5% mark, while their English lexicon reached 13.2% of closed-class items. That shows that they basically had reached the maximum ceilings, even more so in French than in English.

All of the other categories follow a very similar pattern if not parallel. The most striking example is that of the social words in figure 5. The two lines are almost exactly the same for the two languages. The closed-class items represent about 10% of the total vocabulary at the lowest point in both English and French and about 60% at the highest.
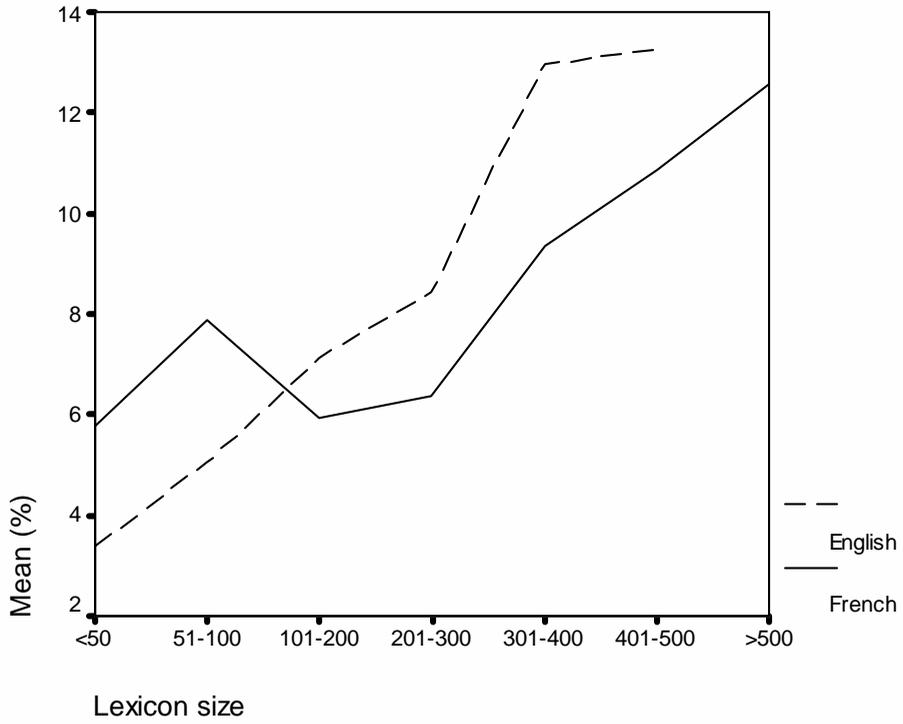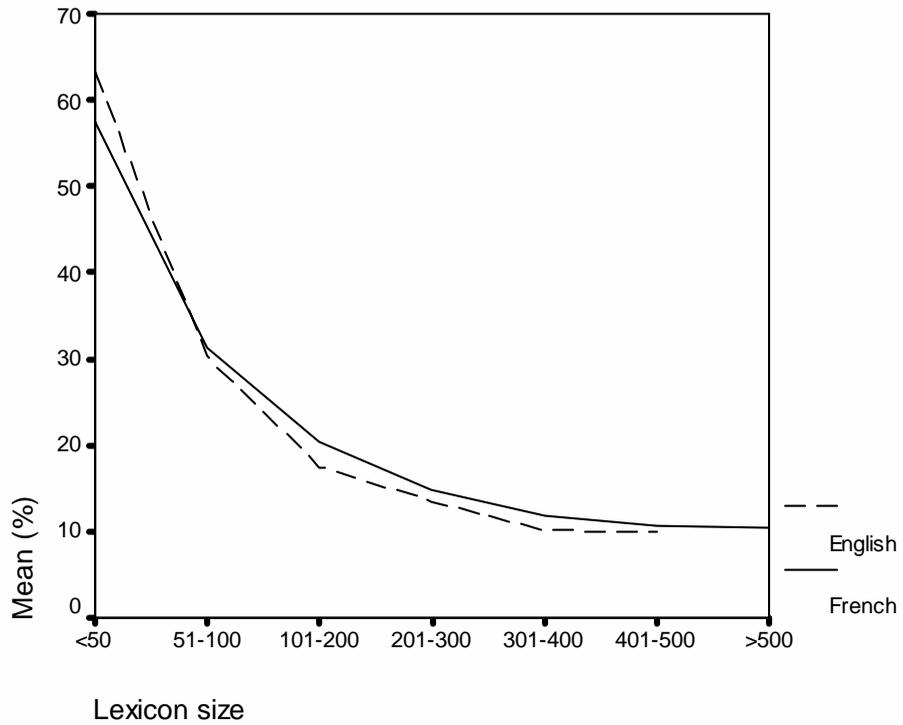
Figure 4. Closed-class items

Figure 5. Social words

Finally, the common nous category also follows a parallel pattern. However, the English category seems to decrease more rapidly than the French one. By the 400-500-word stage, the children had reached the ceiling proportion of nouns in their English lexicon (45%). While, the French lexicon had not yet caught up by the 500-word stage (ceiling was 47.6%, average was 49.3%).

Figure 6. Common nouns
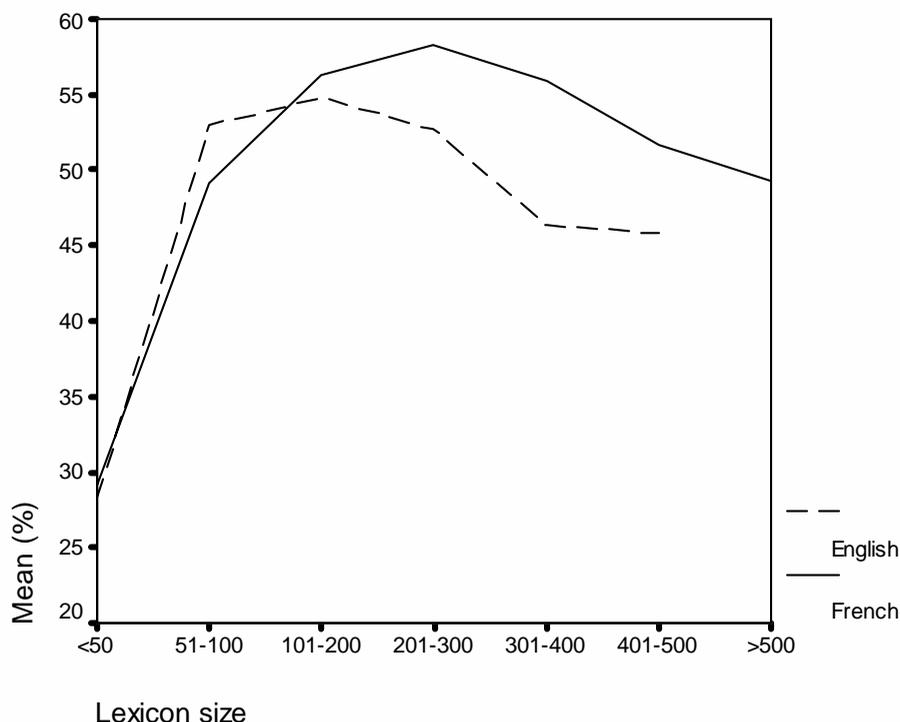
We should now take a more detailed example of a child whose lexicon is 'unbalanced'. By this, we mean that ANT has a much larger vocabulary in French than in English. Despite living in England, she had, at the time of the study, little contact with English speakers. She spent most of her days at home wit her French-speaking mother. By age 2;0, her productive vocabulary was composed of 310 and 53 words, respectively in French and in English. As table 4 shows, ANT is clearly at stage 1 of lexical development in English but already reaching stage 4 in French. Even though, one language seems clearly dominant, English is still developing according to the stages expected as described above.

Table 4. ANT (2;0)

	French	English
Common nouns	181 (58.4%)	36 (67.9%)
Predicates	66 (21.3%)	3 (5.7%)
Closed-class	21 (6.8%)	1 (1.9%)
Social words	38 (12.3%)	12 (22.6%)

5   Discussion

Our results seem, as expected, to indicate very strongly that the bilingual lexicon comprises the same categories as the monolingual one and that these categories evolve exactly at the same speed and in the same proportion in each language whatever the number of words. We can compare our results with those obtained for monolinguals. We will not compare raw numbers as they are neither relevant

nor appropriate, as mentioned earlier. Inside each language, the results are consistent with what was found by other studies. That is to say that, social words appear first and represent the majority of lexical items in early production. The second stage, between 50 and 200 words, sees the rapid increase of nouns. The third stage is characterised by the appearance of predicates and adjectives. And finally during the fourth stage, children acquire grammatical words or closed-class items. We found that stage 3 and 4 probably overlap each other and are not easily separable. So the 4-stage approach proposed by Caselli et al. (1999) and replicated in other languages seems to be valid for bilingual children too. Our first hypothesis can be confirmed: bilingual children go through the same stages of lexical acquisition as monolingual children.

Our second hypothesis was about a possible parallel development in both languages. It appears that each category follows the same distribution and evolution in the two languages or at least for French and English in this context. No major difference has been found. No language seems to be able to hold all of one category while the other one would have none. Even unbalanced, both languages in bilingual children seem to develop following the 4 stages of 'normal' lexical development. However, the children studied here are growing up in a fairly balanced environment, linguistically speaking. The situation might be different for children who acquire one language from home and the second one from their environment. It would be interesting to replicate a similar study in other contexts.

We should now turn to our third aim which was to check for cross-linguistic differences. Regarding the possible larger number of closed-class items in French than in English, we did not find conclusive evidence for that. As Caselli et al. (1999) we did not find that children in either language ever had more than 15% of closed-class items in their lexicon. So Bassano's (1998) 20% seem to have been an isolated case. We did not find either that the French lexicon contained more closed-class items than the English one for our bilingual children contrary to what had been found by Caselli et al (1999). One reason for that could be that the French CDI form actually put ma/mon/mes which is the first person possessive article respectively for singular feminine, singular masculine, and plural all as one item to be ticked by parents. This would automatically lower the scores. However, like Caselli et al. (1999) who found that Italian children acquired closed-class items quicker than English ones, the bilingual children in this study seemed to have more closed-class items in French before English ones in their lexicon. At the first stage (below 50 words), bilingual children had 5.7% and 3.3% of closed-class items respectively for the FCDI and ECDI. This might be due to the fact that there are more closed-class items in French and so children are more likely to pick up one or two earlier on. The way closed-class items are represented on the FCDI, compared with the ECDI, does not reflect the larger number of closed-class items in the French language. So, even if earlier on, children seem to have more closed-class items, the difference might not be maintained simply because of the design of the forms.

Finally, looking at comparisons with monolinguals, we found some indication that predicates might appear earlier in French. We did find that the French lexicon had a larger proportion of verbs earlier on based on the CDIs' data. But the difference was not as significant as in Boysson-Bardies' (1999) study, for example. The diaries revealed that children tend to have one or two verbs in their very early lexicon in French but not in English. The bilingual children seemed to have reached the ceilings of the FCDI by the end of the data collection but not those of the ECDI which might influence the results. It still remains to verify whether this is just the case for these bilingual children or whether the ceilings of the FCDI are also more rapidly reached by monolingual children. Unfortunately, the French study is still in progress. Here again, the design of the forms might be the cause for the difference. This would need to be further investigated.

6   Conclusions

This paper reports only some of the findings of a longitudinal study of 13 French-English bilingual children between 1;0 and 3;0. The findings of this study highlight the fact that bilingual lexical development is in many ways similar to monolingual lexical development. Bilingual children seem to go through exactly the same four lexical stages that had been highlighted cross-linguistically for monolingual children. This is not an expected result as there was no reason to believe that bilingual children would be different. However, what is more interesting is to notice that the two languages do evolve in parallel. None of the children studied had one language that contained words from a category

that was not present in the other language. Even if some of the children were clearly dominant (lexically speaking) in one language, that did not happen. Finally, we found evidence that would tend to show that bilingual children acquire predicates earlier in French than in English. This needs to be further investigated though, as we do not have all of the norms for the different CDIs available yet.

Clearly, further investigations are needed in order to understand the relationship that these results could have on the emergence of grammar and more especially on the emergence of code-switching or mixing. More studies are also required on different language pairs, and especially languages that are structurally more different. We also need to consider bilingual children growing up in different situations than the one-parent-one language strategy. Their lexicon might then show different developmental patterns.

References

Bassano, Dominique (2000). Early development of nouns and verbs in French: exploring the interface between lexicon and grammar. *Journal of Child Language, 27*, 521-559.

Bassano, Dominique, Maillochon, I., & Eme, E. (1998). Developmental changes and variability in the early lexicon: a study of French children's naturalistic productions. *Journal of Child Language, 25,* 493-531.

Bates, Elizabeth, Marchman, V., Thal, D., Fenson, L., Dale, P. S., Reznick, J. S., Reilly, J., & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language, 21*, 85-123.

Benedict, Helen (1978). Early lexical development: comprehension and production. *Journal of Child Language, 6*, 183-200.

Boysson-Bardies, Bénédicte (1999). *How language comes to children: from birth to two years*. Cambridge, MA.: MIT Press.

Caselli, Cristina, Casadio, P., & Bates, E. (1999). A comparison of the transition from first words to grammar in English and Italian. *Journal of Child Language, 26*, 69-111.

Dale, Philip S., Bates, E., Reznick, J. S., & Morisset, C. (1989). The validity of a parent report instrument of child language at twenty months. *Journal of Child Language, 16*, 239-249.

Deuchar, Margaret, & Quay, S. (2000). *Bilingual Acquisition (Theoretical implications of a case study)*. Oxford: Oxford University Press.

Dromi, Esther (1996). *Early lexical development*. San Diego: Singular Publishing Group.

Eriksson, Mårten (2001). *Children's words in inventories and tests: what questions do they answer?* Paper presented at the ELA 2001, Lyon, France.

Fenson, Larry, Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J. (1993). *The MacArthur Communicative Development Inventories: User's Guide and Technical Manual*. San Diego: Singular Publishing Group.

Harrison, Claire (2001). *Validity and preliminary norms for a British English adaptation of the MacArthur Communicative Development Inventories: Toddlers*. Unpublished thesis, University of Newcastle-upon-Tyne.

Jackson-Maldonado, Donna, Thal, D., Marchman, V., Bates, E., & Gutierrez-Clellen, V. (1993). Early lexical development in Spanish-speaking infants and toddlers. *Journal of Child Language, 20*, 523-549.

Jones, Michael A. (1996). *Foundations of French syntax*. Cambridge: Cambridge University Press.

Kern, Sophie (2001). *Développement des premiers comportements linguistiques chez le jeune enfant* (Rapport d'activités APN 98-2000).

Kern, Sophie (submitted). Lexicon development in French-speaking infants. *Applied Psycholinguistics*.

Klee, Thomas, & Harrison, C. (2001). *CDI Words and Sentences: validity and preliminary norms for British English.* Paper presented at the Child Language Seminar, University of Hertfordshire.

Maital, Sharone L., Dromi, E., Sagi, A., & Bornstein, M. H. (2000). The Hebrew Communicative Development Inventory: language specific properties and cross-linguistic generalizations. *Journal of Child Language, 27*, 43-67.

Marr, Claire (1999). *The concurrent validity of British English adaptations of the long and short forms of the MacArthur Communicative Development Inventories (toddlers)*. Unpublished thesis, University of Newcastle-upon-Tyne.

Pearson, Barbara Z. (1998). Assessing lexical development in bilingual babies and toddlers. *International Journal of Bilingualism, 2*(3), 347-372.

Pearson, Barbara Z., Fernandez, S., & Oller, D. K. (1995). Cross-language synonyms in the lexicon of bilingual infants: one language or two? *Journal of Child Language, 22,* 345-368.

Robertson, Elaine (1999). *The effect of list length on parent-report measure of children's early vocabulary production.* Unpublished thesis, University of Newcastle-upon-Tyne.

ISB4: Proceedings of the 4th International Symposium on Bilingualism

edited by James Cohen, Kara T. McAlister, Kellie Rolstad, and Jeff MacSwan

Cascadilla Press   Somerville, MA   2005

Web access and citation information

This paper is available from www.cascadilla.com/isb4.html and is identical
to the version published by Cascadilla Press on CD-ROM and in library binding.