ARTICLES

Assessing Language Dominance in Bilingual Acquisition: A Case for Mean Length Utterance Differentials

Virginia Yip
Chinese University of Hong Kong

Stephen Matthews
University of Hong Kong

The notion of language dominance is often defined in terms of proficiency. We distinguish dominance, as a property of the bilingual mind and a concept of language knowledge, from proficiency, as a concept of language use. We discuss ways in which language dominance may be assessed, with a focus on measures of mean length of utterance (MLU). Comparison of MLU in the child's 2 languages is subject to questions of comparability across languages. Using the Hong Kong Bilingual corpus of Cantonese–English children's development, we show how MLU differentials can be a viable measure of dominance that captures asymmetrical development where there is an imbalance between the child's 2 languages. The directionality of syntactic transfer goes primarily from the language with higher MLU value to the language with lower MLU value, and the MLU differential matches the pervasiveness of transfer effects, as in the case of null objects discussed here: The greater the differential, the more frequent the occurrence of null objects. Cantonese-dominant children with a larger MLU differential use null objects more frequently than those with a lower MLU differential. In our case studies, MLU differentials also matched with language preferences and silent periods but did not predict the directionality of code-mixing.

Correspondence should be addressed to Virginia Yip, Department of Linguistics and Modern Languages, Chinese University of Hong Kong, Shatin, N. T., Hong Kong. E-mail: vyip@humanum. arts.cuhk.edu.hk

This article discusses some issues concerning the assessment of language dominance in the early stages of child bilingual acquisition. The notion of language dominance has often been invoked in research on bilingual first-language acquisition, though without wide consensus on how it should be defined and measured over time. On the basis of the Hong Kong Bilingual Child Language Corpus, we discuss how researchers have defined and measured language dominance, raising the issue with regard to the distinction between dominance as a property of the bilingual mind and proficiency as a concept of language use. We argue that language dominance must be related to underlying competence (i.e., knowledge of language) to be compatible with a competence-based theory of transfer. Our discussion centers on the mean length of utterance (MLU) of the two languages in the bilingual children as a measure of language dominance. We argue that despite a number of long-standing unresolved problems, MLU differentials can be a useful quantifiable objective measure of dominance.

BILINGUAL INPUT AND LANGUAGE DOMINANCE

Compared with acquiring a single language in a monolingual context, the acquisition of two languages in bilingual or multilingual contexts poses even more challenges to the child on many grounds, beginning with the quantity of input: Bilingual children have their input space divided, so the amount and frequency of exposure to each language at any given time is smaller than that of monolinguals acquiring each language (Paradis & Genesee, 1996). When the input is less than balanced, one of the two languages may develop faster or show greater complexity at a given age. This language is said to be dominant. How language dominance is measured at a given stage of development and over time is a complex issue that is still under-researched.

THE HONG KONG BILINGUAL CHILD LANGUAGE CORPUS

The data for this study are drawn from the Hong Kong Bilingual Child Language Corpus, developed by the authors and available through the Child Language Data Exchange System (MacWhinney, 2000; see http://childes.psy.cmu.edu/data/Biling). The data released to date contain longitudinal data for 6 bilingual Cantonese–English participants ranging from 1;03 to 4;06. There are a total of 430 tagged files in two languages. The participants were exposed to Cantonese and English from birth and grew up in one parent–one language families in Hong Kong in which each parent is a native speaker of Cantonese or English. The background information of the participants is summarized in Table 1.

TABLE 1
Background of Bilingual Participants

Participants	Native Language of Mother	Native Language of Father	Age Span of Study
Timmy	Cantonese	English	1;05.20–3;06.25
Sophie	Cantonese	English	1;06.00–4;00.00
Kathryn	English	Cantonese	2;09.23–4;06.07
Llywelyn	Cantonese	English	1;06.00–3;05.28
Charlotte	Cantonese	English	1;05.10–3;06.14
Alicia	Cantonese	English	1;03.10–3;00.24

In each recording session, two research assistants took turns interacting with the participant, each speaking and eliciting a different target language for 30 min. Parents or siblings were sometimes present. The interaction was audiotaped and, in a majority of cases, videotaped as well. Each transcript consisted of approximately one half an hour of naturalistic interaction between the child and the adult interlocutors in the participant's home. For 3 participants (Kathryn, Llywelyn, and Charlotte), the transcripts were linked to digital audio files, whereas for 1 participant (Alicia), the full set of transcripts were linked to video files so the action could be viewed alongside the transcript. The introduction of browsable XML format made the data even more accessible, because the data could be viewed online without the need to download them. This corpus is now the largest multimedia bilingual child language corpus in the Child Language Data Exchange System archive.

Our discussion focuses on the 3 siblings (Timmy, Sophie, and Alicia) on whom the most work has been done to date. As they grew up in the same family, the input conditions are similar, and the corpus data are supplemented by diary data. Reference to the other 3 children, on whom work is still ongoing at the time of writing, is made where relevant.

LANGUAGE DOMINANCE IN EARLY BILINGUAL DEVELOPMENT

In the bilingual acquisition literature, the notion of language dominance is used, pretheoretically, to describe a situation in which one of a child's languages is more advanced or developing faster than the other.[1] Dominance is argued to be important in identifying types and degrees of bilingualism (Romaine, 1995). The importance of dominance extends into later childhood and adulthood, where it is ac-

---

[1]It may be significant that there is no established term for the child's "other" language(s). In genetics, the counterpart of "dominant" is "recessive," but this term does not seem appropriate in bilingual development, because recessive genes are typically unexpressed (rather than under-expressed). Bernardini and Schlyter (2004) used the term "nondominant" before introducing their own distinction between "stronger" and "weaker" languages.

knowledged that an individual child's pattern of dominance can change with shifting home and school environments. Thus Baker and Prys Jones (1998) observed that "in the majority of bilinguals one language is more dominant than the other" (p. 12). One point that is uncontroversial is that dominance is by no means static: Dominance patterns may change over time depending on individual experiences (Romaine, 1995). The dynamic nature of dominance is illustrated later with our bilingual subjects' corpus data. Furthermore, dominance may also have a role in predicting and explaining outcomes. For example, certain nontarget features in a bilingual child's nondominant language tend to persist or are unlearned slowly over a protracted period of time (cf. Yip & Matthews, 2005) or aspects of the nondominant language may fail to develop fully or be lost (Romaine, 1995).

Defining Language Dominance

The term *dominance* is often defined in terms of proficiency: The dominant language is that "in which the bilingual is informally considered to be most proficient" (Petersen, 1988, p. 487; see also Deuchar & Muntz, 2003, and Genesee, Nicoladis, & Paradis, 1995, among others). De Houwer (1998) pointed out that "proficiency" is scarcely used in monolingual child language research, yet it is readily used in a relative sense when two or more languages are involved, as in bilingual child language acquisition. Such invocation of proficiency is "quite at odds with the general thinking in monolingual child language research today" (De Houwer, 1998, p. 259), leading the author to question whether the notion of dominance is needed at all in connection with young bilingual children.

We adopt the view that language acquisition research should be concerned with investigating a learner's knowledge that underlies language use and proficiency (Lakshmanan, 1995). Knowledge of language is taken to be a mentally represented grammar (Chomsky, 1986). For example, when we claim that bilingual children use null objects (see the Language Dominance and Null Objects section), we do not merely mean that learners "omit" objects. Rather, a null object is part of the structure assigned to sentences in the children's grammar. Consider a typical example from diary data as shown in (1):

(1) You get, I eat [father takes chocolates off shelf] (Timmy, 2;02.03)

The child's hypothetical representation for this sentence includes a null topic ("chocolate," the focus of his attention) and two null objects whose reference is determined by the null topic, as formalized in (2).[2]

---

[2]In particular models of grammar, technical distinctions are made between various null elements; thus the null pronominal element might be treated as a variable (x) or a null pronoun (*pro*). Such distinctions are highly theory-dependent, whereas the observations we wish to make transcend such distinctions. In particular, the null pronominal element, however represented, is essential to the interpretation of the sentence. The subscript i indicates that the null objects (x) refer to the same entity as the sentence topic.

(2) [TOPIC$_i$] you get [x_i], I eat [x_i]

In such a view, to be of theoretical interest (e.g., in accounting for how knowledge of language is acquired on the basis of dual input, how two linguistic systems interact in transfer, code-mixing, and so on), dominance must be related to underlying competence and not merely a measure of performance or language use. We assume that syntactic transfer, for example, takes place at the level of competence. If dominance is taken to be merely a performance notion, then it cannot explain the occurrence of transfer, unless one resorts to a surface-oriented and performance-driven concept of transfer. To the extent that the theory of transfer is competence based, the conceptualization of language dominance should be correspondingly competence based for the two notions to be theoretically compatible and coherent.

A view of dominance compatible with these assumptions is that "Language dominance is essentially a psycholinguistic phenomenon closely intermeshed with sociolinguistic parameters" (Lanza, 2004, pp. 172–173) As a psycholinguistic phenomenon, language dominance should be characterized as a property of the mind, albeit influenced by sociolinguistic parameters such as quantity of input and influencing aspects of performance such as fluency. Like Lanza, we believe that the problems that arise in defining dominance do not invalidate the pretheoretical notion.

Language Dominance and Syntactic Transfer

Language dominance is widely considered to be one of the important factors in accounting for the direction of transfer. In our Cantonese–English bilingual participants, transfer is asymmetrical in a way that is predicted by dominance: Transfer from Cantonese to English is clearly visible in many areas of grammar, whereas influence of English on Cantonese, if any, is subtle and much more difficult to demonstrate (Yip & Matthews, 2000). For example, prenominal relative clauses are transferred from Cantonese to English as in (3):

(3) Where's the Santa Claus give me the gun? (Timmy, 2;07.05)
    [i.e., where is the gun Santa Claus gave me?]

This pattern is transferred from the Cantonese relative clause as seen in (4), where the same child is talking about the very same topic:[3]

(4) Santa Claus bei2 lei5 go3 coeng1 le1?
    Santa Claus give you CL   gun   SFP
    "Where's the gun Santa Claus gave me?" (Timmy, 2;8.25)

---

[3]At this stage pronoun reversal is often found in Timmy's Cantonese, so that *lei5* "you" in (4) refers to himself.

The structure transferred in (3) is shown in (5), in which the clause *Santa Claus give me* modifies the noun phrase *the gun*:

(5) Where's [the [Santa Claus give me $_{\text{Rel Cl}}$] the gun $_{\text{NP}}$]?

As discussed by Matthews and Yip (2003), transfer in this case could be attributed either to overall dominance of Cantonese (meaning that the child's Cantonese as a whole is more fully developed than his English during the period when transfer occurs) or to a specific developmental asynchrony between Cantonese and English (such that relative clauses of this type typically develop earlier in Cantonese than in English and undergo transfer in the interim period). Input ambiguity as defined by Müller (1998) does not appear to offer a plausible explanation, because in the English input relative clauses uniformly follow the noun.[4]

The degree of language dominance may also play a role in determining the pervasiveness of transfer effects in certain domains of grammar. In the section titled Language Dominance and Null Objects, we show that dominance differentials predict differential rates of transfer in the domain of null objects.

Measuring Dominance

Objective measures of development include several measures adopted from research on first-language acquisition (Brown, 1973). Those used by Genesee et al. (1995) and Deuchar and Muntz (2003) are

- MLU.
- upper bound (length of the longest utterance in a given sample).
- multiword (or multimorpheme) utterances (percentage of utterances containing more than one word or morpheme).
- word types (or verb types): number of different lexical items used in a sample.

To apply these measures to dominance in bilingual children, the measures must be comparable across the two (or more) languages being acquired: "One must have a baseline for cross-linguistic comparison that works equally well for both languages. This is a fundamental problem that has so far not been solved" (De

---

[4]The prenominal relative clause as used by the children is, however, consistent with the prenominal placement of adjectives in English. This, together with the continuum between adjectival and clausal modifiers in Cantonese (Matthews & Yip, 2003) may provide a motivation for prenominal relatives in Cantonese–English bilingual children. We thank Nik Gisborne and Usha Lakshmanan for discussion of this point.

Houwer, 1998, p. 258) These problems are discussed in the Measuring Dominance With MLU section in relation to MLU.

Another set of criteria is specific to bilingual first-language acquisition:

- Language preference: In some situations (such as recording sessions) children are reluctant to use a certain language. If this behavior is systematic, the language that the child is more willing to speak is considered to be dominant (Saunders, 1988).
- Direction of language mixing: Here the proposal is that when speaking his or her weaker language, a child is more likely to resort to words from the stronger language, resulting in code-mixing or switching. Swain and Wesche (1975) found French functional morphemes mixed with English in their participants, attributing this to French being the stronger language. A recent formulation of this idea is the Ivy Hypothesis of Bernardini and Schlyter (2004), according to which functional elements of the "stronger language" are retained when speaking the "weaker language." Lanza (2004) also invoked directionality of mixing as evidence of language dominance.

We discuss the relation between these phenomena and dominance in the Other Indicators of Dominance section.

MEASURING DOMINANCE WITH MLU

As observed by Döpke (1998), MLU is useful for within-language comparisons but may not be directly comparable across languages, especially those of different morphological types. If a child is acquiring an agglutinating language such as Turkish together with an isolating language such as Cantonese, for example, MLU measured in words (MLUw) will not be comparable because of the greater complexity of individual words in Turkish. In an agglutinating language, numerous affixes may be attached to a word stem. Consequently, at the same stage of development, we can expect a higher MLU in Cantonese than in Turkish. In such cases, a possible solution is to measure MLU in terms of morphemes (MLUm) rather than words (MLUw). If the corpus is transcribed in such as way as to mark morpheme divisions, MLU (as computed automatically by CLAN software) can be counted in terms of morphemes per utterance, and the resulting figures (MLUm) will provide a more comparable measure of complexity. This option, however, is only feasible if (a) both languages allow segmentation of morphemes (which fusional languages such as Russian generally do not) and (b) this segmentation is fully coded in the available corpus.

These issues arose in a study of Swedish–French and Swedish–Italian bilingual children by Bernardini and Schlyter (2004) who used MLU measured in MLUw to

demonstrate that the children had "stronger" and "weaker" languages. The authors noted that Swedish is expected to have lower MLUw values than French or Italian, as it has additional bound morphemes. For example, Swedish has a suffixed definite article, as in *stad-en* "the city" whereas Italian has a separate word (*la città* "the city"). The same expression "the city" is therefore counted as one word in Swedish and two words in Italian. Despite such "deflationary" factors, the participants in fact showed a higher MLU in Swedish. It is therefore most unlikely that the discrepancy between the MLUw measures for the children's two languages is an artifact of morphological differences between the languages being acquired; indeed, the true discrepancy is even stronger than it appears in their MLU charts (Bernardini & Schlyter, 2004).

The calculation of MLUw depends on decisions regarding what constitutes a word—a problem that has not been resolved, either in general (see Dixon & Aikhenvald, 2002) or specifically with regard to Chinese (see Packard, 2000). In particular, the concepts of phonological word, morphological word, and syntactic word do not necessarily match. Our MLUw calculations are based on the word divisions as made in the transcripts of the Hong Kong Bilingual Child Language Corpus. The transcription is modeled on the description of Cantonese grammar by Matthews and Yip (1994) except that (for consistency with the Codes for the Human Analysis of Transcripts format) the hyphen notation is not used to show word-internal divisions.[5] Consequently, morphemes such as aspect markers are transcribed as separate words, arguably "inflating" the MLU for Cantonese.

Yip and Matthews (2000) suggested two lines of response to this problem. First, Cantonese and child English can both be treated as predominantly isolating languages, especially because in the participants' English, inflectional, morphology is not yet in place. In an ideal isolating language, by definition, each morpheme is a separate word.[6] Cantonese is not a perfect isolating language in this sense as it has compound words composed of two or more morphemes as in (6) and a small number of bound affixes, some of which are frequent in child language, such as the perfective aspect marker *zo2* in (7):

(6) fan3-gaau3
    lie-asleep
    "sleep"

(7) sik6-zo2
    eat-PFV
    "have eaten"

---

[5]The transcription conventions given in the Codes for the Human Analysis of Transcripts manual are available at http://childes.psy.cmu.edu/.

[6]In the isolating type, "Affixes are absent" and "there are monosyllabic words, both lexical … and grammatical" (Sgall, 1995, p. 56).

Klee, Stokes, Wong, Fletcher, and Gavin (2004) treated compounds such as (6) as single words but aspect markers such as *zo2* in (7) separately.[7] By transcribing both these items as separate words, it is likely that our transcripts inflate the MLU for Cantonese. Another relevant feature of Cantonese is sentence-final particles. Although these are generally regarded as separate words, it is arguable that their inclusion inflates MLU, because even at what is essentially the one-word stage, children often add a particle, as seen in (8):

(8) Child: tung-tung aa3
    hurt-hurt   SFP
    "It hurts."
    Mother: lei5 tung3 aa4?
    You hurt   SFP
    "You're hurting?"
    Child:ni1dou6 aa3
    here   SFP
    "Here." (Alicia, 1;04.12)

Other factors may compensate for these inflationary factors, however. One such factor is the presence of null subjects and objects in Cantonese, as noted by Klee et al. (2004). As discussed in relation to (1) previously, we consider these null elements to be part of the structure of the sentence, but they are disregarded in the computation of MLU, whereas the corresponding English sentences would contain overt subjects and objects. MLU therefore underestimates the complexity of Cantonese utterances containing null subjects and objects. In establishing a baseline for monolingual children, Klee et al. showed that MLU rises more sharply with age in English (probably as a result of acquisition of inflectional morphology) than in Cantonese; however, between 30 and 36 months the MLU curves for the two languages intersect. At this stage, the MLU values should therefore be closely comparable.

Even if we conclude that MLU values are not strictly comparable across the two languages, it is still possible to use MLU differentials between a bilingual's two languages to compare individual bilingual children with each other and to chart changes in dominance patterns over time. Consider Figure 1 showing the development of MLUw in Timmy's two languages.

Figure 1 appears to show that Cantonese developed faster than English in the period 2;01 to 2;08, whereas after age 2;09 the MLU values were closely

---

[7]A potential problem arises when compounds as in (6) are treated as single words but aspect markers as in (7) are treated as separate words, as in Klee et al. (2004). The two cases in (6) and (7) are related because aspect markers come between the two parts of the compound verb, as in *fan3 zo2 gaau3* "has gone to sleep." Consequently, if aspect markers are treated as separate words, it is logical to separate the parts of the compound verb too.
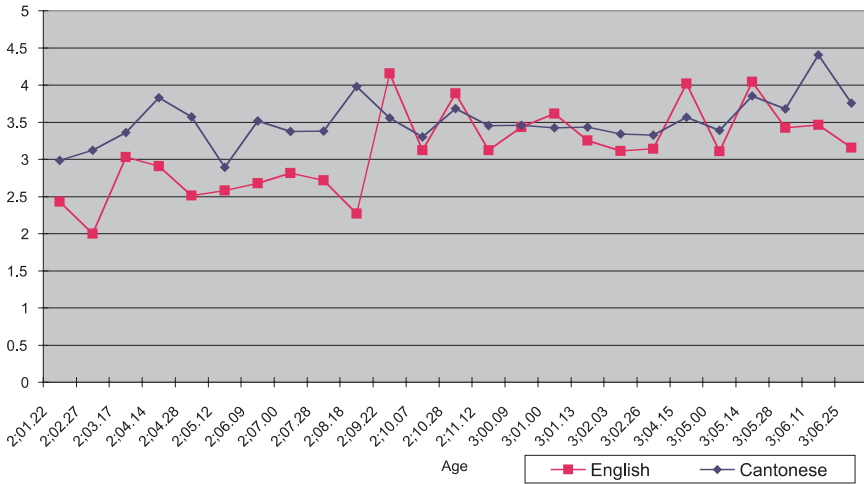
FIGURE 1   MLU for Timmy's Cantonese and English.

matched.[8] Given the uncertainty concerning comparability of MLUw across lan-
guages, however, this pattern allows for a number of interpretations. To the extent
that the measures for the two languages concerned are comparable, the pattern
shown is one of a period of Cantonese dominance followed by a period of more
balanced development. If the first period is an artifact of the calculation method
(e.g., if the word divisions assumed for Cantonese have the effect of substantially
inflating MLUw), then the Cantonese MLU should be lowered across the board:
the gap between English and Cantonese would close in the initial period, followed
by a period of English dominance after age 2;09. Alternatively, if the Cantonese
MLU is somehow underestimated, the evidence for dominance is even stronger
than Figure 1 suggests. Because there is no independent reason to assume English
dominance (given the child's input conditions and language preferences), the first
interpretation seems most plausible. In the period 2;01 to 2;08 Cantonese is domi-
nant relative to the later period of development, where the two languages are more
evenly balanced.

Timmy's two younger sisters, Sophie and Alicia, show a more consistent pat-
tern. Throughout the period of study, the MLU is higher for Cantonese than for
English (see Figures 2 and 3).

---

[8]The MLU values and differentials for Timmy are computed only for those files for which both
English and Cantonese files are available from the same date, giving a total of 25 sampling points. The
total corpus contains 34 Cantonese and 38 English transcripts, all of which are used to calculate the rate
of null objects in English as shown in Table 3. No such discrepancy arises in the data sets for Sophie and
Alicia, in which matching Cantonese and English files are available at each sampling point.

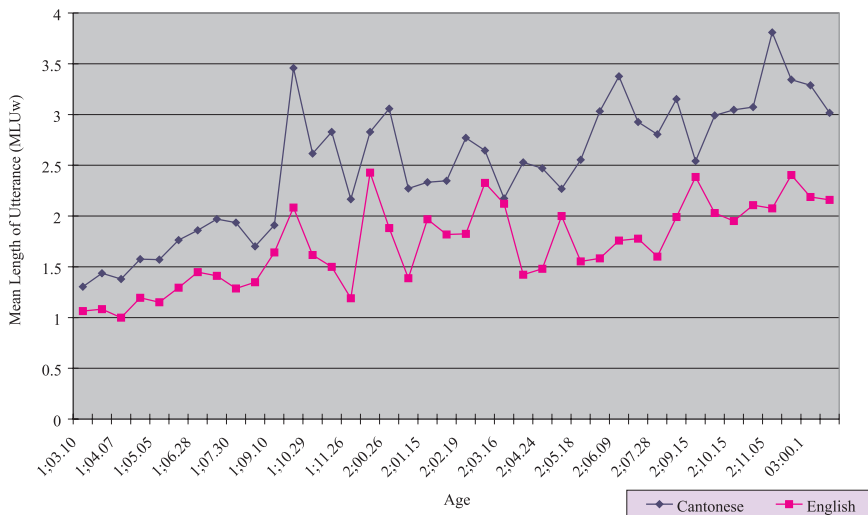FIGURE 2   MLU for Sophie's Cantonese and English.

FIGURE 3   MLU for Alicia's Cantonese and English.

Over the entire period of study, the mean Cantonese MLU values for Sophie and Alicia were nearly one MLU point above those for English (Table 2). We do not believe that the difference can be merely an artifact of problems of comparability in MLUs for English and Cantonese, because (a) the differences are too high to be attributed to the inflationary factors discussed previously and (b) the pattern relates clearly to the children's observed language production; between ages 1 and 2, for example, Sophie understood English but produced only occasional words, whereas she was producing whole sentences in Cantonese (also see the following section). We therefore proposed to use MLU differential, as defined in (9), as a measure of language dominance.[9]

(9) MLU differential: The difference between MLU scores for a child's two languages at a given sampling point or (expressed as a mean) over a period of development.

The dynamic nature of dominance is evident in that the fluctuations in MLU values in two languages and their differentials can be captured at a particular time interval as well as over an extended period of time. Table 2 shows MLUw differentials for the 3 siblings represented in the corpus.

A further methodological issue concerns how to measure MLU when code-mixing occurs in the corpus. This could be another inflationary factor (e.g., English words occurring in Cantonese data files would inflate the MLU figures for Cantonese). Bernardini and Schlyter (2004) counted only "pure" utterances, excluding those containing code-mixing for the purpose of MLU calculation. In the corpus for Timmy, Sophie, and Alicia, we used a post-code to mark code-mixed words, which were then ignored by CLAN in calculating MLU.[10]

Language Dominance and Null Objects

Pursuing the idea of dominance as measured by MLU differentials, based on Table 2 it appears that Sophie and Alicia (with mean MLU differentials of 0.85 and 0.79, respectively) were more strongly and consistently Cantonese-dominant than Timmy, who showed a differential of 0.39.[11] We believe that this contrast is borne

---

[9]Bernardini & Schlyter (2004) effectively use the same concept when they note that "uneven development" in their participants is characterized by a difference of about one MLU point between the child's two languages.

[10]As the marking of post-code has yet to be implemented for the remaining 3 children (Kathryn, Llywelyn, and Charlotte), comparable MLU figures are not available at this time.

[11]The mean MLU values for both languages were higher in the case of Timmy because the corpus covered a somewhat later period of development, extending from 2;01.22 to 3;06.25. Prior to age 2;01 the recordings did not separate the languages, resulting in mixed files (which are also available at the Child Language Data Exchange System archive).

TABLE 2
Mean MLU and MLU Differential in Three Siblings

	Timmy[a]	Sophie[b]	Alicia[c]
English	3.12	1.73	1.71
Cantonese	3.51	2.58	2.50
MLU differential	0.39	0.85	0.79

Note.   MLU = mean length of utterance.
[a]25 files per language from 2;01.22–3;03.25. [b]40 files per language from 1;06.00–3;00.09. [c]40 files per language from 1;03.10–3;00.24.

out when we examine the relation between dominance and transfer. One of several properties of Cantonese that is systematically transferred in our participants' English involves null objects. A typical example illustrating the null object phenomenon involves the verb *put* as in (10):

(10) Investigator: Where shall we stick it?
    Child: Put here. (Timmy, 2;05.05)

This nontarget-like structure whereby *put* is directly followed by a locative resembles the Cantonese structure as seen in (11), in which both the adult and the child use the verb *baai2* "put" with a null object.

(11) Investigator: Sik1   m4   sik1 baai2 aa3?
                know   not   know put   SFP
    "Do you know how to put it in place?"
    Child: Baai2   aa1
          put   SFP
    "Put it in place." (Sophie, 1;06)

Yip and Matthews (2000, 2005) showed that the bilingual participants' use of null objects is quantitatively and qualitatively different from that of monolingual children in ways that are consistent with transfer from Cantonese in this domain.

    As shown in Table 3, the overall frequency of null objects for the five verbs studied was higher in Sophie and Alicia's English than in Timmy's. Among the most frequent verbs that took an obligatory object in English, the verb *put* stands out as taking a null object 36% of the time in Timmy's data between 2;00.26 and 3;06.25, compared to 59% in Alicia's English from 1;03.10 to 3;00.24 and 82% of occurrences in Sophie's data between 1;06.01 and 3;00.09.

    That *put* should be the single verb most frequently used with a null object no doubt relates to its argument structure: As a verb that uniquely requires a prepo-

TABLE 3
Frequency of Null Objects With Five Transitive Verbs
in the Bilingual Corpus[a]

	No. of Null Objects			Total No. of Occurrences of Each Transitive Verb			Null Objects (%)		
Verb	Timmy	Sophie	Alicia	Timmy	Sophie	Alicia	Timmy	Sophie	Alicia
Get	6	2	0	122	7	5	5	29	0
Like	3	16	8	88	60	62	3	27	13
Put	34	18	37	95	22	63	36	82	59
Take	10	19	1	41	32	4	24	59	25
Want	6	45	54	67	184	160	9	24	34
Total	59	100	100	413	305	294	14	33	34

[a]Timmy: 38 files from 2;00.26–3;06.25; Sophie: 40 files from 1;06.01;–3;00.09; Alicia: 40 files from 1;03.10–3;00.24.

sitional phrase as well as a direct object, it calls for a relatively complex and heavy verb phrase of the form [V NP PP] if all the arguments are to be overtly realized.[12] Omitting the object therefore simplifies the verb phrase to a [V PP] configuration. Other verbs including *get, like, take,* and *want* varied in their frequency of taking a null object, ranging from 3% to 59%, but for each verb the frequency of null objects was at least twice as high in Sophie's English as in Timmy's. Overall, for the five verbs studied, Sophie and Alicia had a null object rate of 33% and 34%, respectively, compared to 14% for Timmy. Thus a quantitative aspect of syntactic transfer, namely frequency of null objects (as seen in Table 3), appears to correlate with degrees of dominance as measured by MLU differentials (as given in Table 2).

This account makes several further predictions that should be testable. For example, 2 of the bilingual children in the Hong Kong corpus lack the pattern of Cantonese dominance as measured by MLU: Kathryn appears to be balanced, whereas Charlotte appears to be English dominant for at least part of the period studied (Lai, 2005). Although these participants are expected to produce some null objects in their English (just as monolingual children do), we predict that they will show significantly lower percentages of null objects than Cantonese-dominant children such as Timmy, Sophie, and Alicia. Ongoing work by Emily Yiu (2005) has suggested that the null object rate for Kathryn is in fact under 10%.

---

[12]A reviewer pointed out exceptional collocations such as the athletics term *put the shot*, in which the PP argument is obligatorily omitted. We assume that such domain-specific terms are not instantiated in the input to the children and can safely be ignored for this argument.

OTHER INDICATORS OF DOMINANCE

As noted previously, other indicators of dominance include language preferences and directionality of mixing. We now discuss how these indicators relate to the picture of dominance as measured by quantitative measures such as MLU differentials.

Language Preferences and Silent Periods

Language preferences can often be observed in bilingual children's interaction with their interlocutors. In the recording sessions on which the Hong Kong corpus is based, it was necessary for one researcher to interact with the child in English and the other in Cantonese, each for a half-hour period. Frequently the child would be reluctant to use one language or the other. Sometimes this was part of a consistent pattern over a period, as with Charlotte (identified as English-dominant based on her MLUs) who was often reluctant to speak Cantonese in a recording session. These patterns, however, are overlaid with temporary and idiosyncratic ones, such as having spoken largely English on the day of recording. Even Kathryn (considered the most balanced of the participants) showed such preferences in the form of reluctance to speak one language or the other. It is not clear how such temporary factors can be separated from longer term preferences.

A related phenomenon that may be more amenable to systematic study is that of "silent periods." As observed earlier, Sophie went though a silent period between ages 1 and 2 in which she understood English but did not respond or responded with single words only. Alicia went through a similar period of struggling to produce English and occasionally resorted to using Cantonese (sometimes accompanied by an apologetic look) to respond to her father's English. In the early corpus we often saw Sophie responding in Cantonese to questions posed by the investigators in English, as in (12) and (13):

(12) Investigator: What is in this cat's hand?
    Investigator: What is in his hand?
    Child: *baak6tou3*.
            rabbit
            "A rabbit."
    Investigator: *baak6tou3* okay, it's a rabbit. (Sophie, 2;00.18)

(13) Investigator: Sophie, what's this?
    Child: *<jyu2 aa3> [/] jyu2 aa3*.
            fish PRT   fish PRT
    "A fish."
    Investigator: It's a fish?

Investigator: Where's the fish?
Child: Here. (Sophie, 2;02.16)

These observations are consistent with the MLU figures suggesting that the girls' Cantonese was ahead of their English throughout the period of study (see Figures 2 and 3).

The silent period, a time lag between comprehension and production, is generally associated with second-language acquisition (see, e.g., Krashen & Terrell, 1983). It might be argued that children showing a silent period, like Sophie and Alicia, are de facto cases of child second-language acquisition; yet they fit most definitions of bilingual first-language acquisition, given that the child is exposed to both languages from birth. We therefore conclude that a silent period can be a feature of unbalanced bilingual first-language acquisition. It remains to be seen to what extent the development of such children is similar to or different from clear cases of child second-language acquisition, whereby exposure to the second language only begins at age 2 or 3.

Code-Mixing

Another phenomenon thought to be related to language dominance is code-mixing. For example, "a *propensity* to use a certain directionality of mixing can be an indicator of dominance" (Lanza, 2004, p. 173; her emphasis). Although rates of code-mixing can be quantified, we do not see this as a potential measure of dominance; rather, like syntactic transfer as argued in earlier, it would be an aspect of language production that might be partially explained by dominance. Petersen (1988) found that an English–Danish bilingual child used English morphology with Danish lexical items but not vice versa. This asymmetry was attributed to dominance of English; that is, the inflectional morphology of the dominant language is used with lexical morphemes of the nondominant language. A recent formulation of this idea is Bernardini and Schlyter's (2004) Ivy Hypothesis, whereby it is proposed that children will resort to the functional elements of their stronger language when speaking the weaker language. Thus, when Swedish is the stronger language, we find Swedish functional items such as determiners combined with lexical items from the weaker language.

In our participants, asymmetries in directionality of mixing were observed: English items were mixed into Cantonese speech more frequently than vice versa. However, we doubt whether these asymmetries can in general be attributed to dominance. Even Kathryn, who showed the most balanced pattern of development among all the children studied, mixed English items in her Cantonese more than vice versa (Yiu, 2005). Charlotte, a bilingual child showing unbalanced development, was English dominant as indicated by MLU values and language preference. She nevertheless showed bidirectional code-mixing, and her patterns of mixing did

not conform to the predictions of the Ivy Hypothesis as outlined previously (Lai, 2005). To account for these findings, we propose that individual patterns of dominance interact with code-mixing as instantiated in the input to the children. In the Hong Kong context, code-mixing was part of the adult Cantonese input; that is, English words, especially nouns and verbs, are commonly inserted into Cantonese discourse (see, for example, Chan, 2003). This practice is particularly widespread in middle-class families such as those of our participants. Much of the Cantonese–English code-mixing produced by the children was in fact adult-like and may simply show that the child was acquiring code-mixing in response to input of this kind. In the most typical case, an English noun was inserted into a Cantonese utterance, as in example (14) from Yiu (2005) where Kathryn followed the adult investigator's lead in inserting the noun *turtle*:

(14) Investigator: ngo5dei6   waan2   zo2   #   turtle   sin1   laa,   hou2 mou2 aa3
                     We   play   Asp   turtle   first SFP   good not   SFP
      "Let's play the turtle first, shall we?"
      Child: ngo5dei6   jau5   [/]   jau5   loeng5   go3   turtle gaa3.
                     We   have   have   two   CL   turtle SFP
      "We have two turtles!" (Kathryn, 3;02.19)

At most, a subset of cases could be developmental in nature, such as a response to a temporary lexical gap in one language. In the following example, the child inserts the English word *ant* into a Cantonese sentence. The adult prompts her with the Cantonese equivalent, which she then uses herself:

(15) Child: ngo5   zek3   ant   le   1.
              I   CL   ant   PRT
      "Where's my ant?"
      Investigator: Ant?
      Child: ant   hai6   gam2   joeng2   gaa3
              ant   is   like   that   PRT
      "Ants are like that."
      Investigator: ngai5   hai6   mai6   aa3
                       ant   be   not-be   PRT
      "You mean ants?"
      Child: ngo5   waak6   ngai5   sin1   aa2
              I   draw   ant   first   PRT
      "I'll draw an ant first." (Sophie, 2;03.14)

The general implication of this line of work is that code-mixing is a feature of most bilingual societies that is acquired in response to the input. In such environments, it will always be difficult, if not impossible, to separate the acquisition of adult-like

mixing behavior from developmental processes reflecting underlying competence in each language. In speech communities in which code-mixing is not a general feature of the input, this factor is removed and code-mixing may indeed be an essentially developmental feature that could serve as an index of dominance.

CONCLUSIONS

The notion of language dominance is applicable to many bilingual children in whom development is unbalanced. We have argued that dominance should be viewed as a property of the bilingual mind and as distinct from proficiency as a property of language use. Language dominance can be measured objectively using quantitative measures such as MLU. There are well-known problems in comparing MLU values across particular language pairs, especially where different morphological types are involved. Comparability of languages cannot be reliably obtained when the same measure cannot be equally applied to both languages under investigation. Nevertheless, we have shown how MLU differentials between the two languages under investigation can be used to compare children acquiring the same language pair.

There is currently no consensus as to whether dominance as measured by MLU has any deeper significance at the level of language competence apart from its relevance at the level of language use. We have offered evidence that it does, based on syntactic transfer as observed in the Hong Kong Bilingual Child Language Corpus. In our study of 3 siblings, transfer from the dominant language was systematic in certain domains, as in the case of null objects that are transferred from Cantonese to English. Cantonese-dominant children with a larger MLU differential were found to use null objects more frequently than those with a lower MLU differential.

Observational criteria such as language preferences shown by the bilingual child in choosing a particular language in certain language contexts and silent periods can be used to corroborate quantitative measures such as MLU differentials. In the case of Cantonese–English bilingual children, code-mixing is not closely tied to dominance due to the extensive use of code-mixing in the speech community and hence in adult input to our participants.

ACKNOWLEDGMENTS

REFERENCES

Baker, C., & Jones, S. P. (1998). *Encyclopedia of bilingualism and bilingual education.* Clevedon, England: Multilingual Matters.

Bernardini, P., & Schlyter, S. (2004). Growing syntactic structure and code-mixing in the weaker language. *Bilingualism: Language and Cognition, 7,* 49–69.

Brown, R. (1973). *A first language: The early stages.* Cambridge, MA: Harvard University Press.

Chan, B. H.-S. (2003). *Aspects of the syntax, the pragmatics and the production of code- switching.* Frankfurt, Germany: Lang.

Chomsky, N. (1986). *Knowledge of language*. New York: Prager Books.

De Houwer, A. (1998). By way of introduction: Methods in studies of bilingual first language acquisition. *Bilingualism: Language and Cognition, 2,* 249–263.

Deuchar, M., & Muntz, R. (2003). Factors accounting for code-mixing in an early developing bilingual. In Müller Natascha (Ed.), *(In)vulnerable domains in multilingualism* (pp. 161–190). Amsterdam: Benjamins.

Dixon, R. M. W., & Aikhenvald, A. Y. (Eds.). (2002). *Word: A cross-linguistic typology.* Cambridge, England: Cambridge University Press.

Döpke, S. (1998). Competing language structures: The acquisition of verb placement by bilingual German-English children. *Journal of Child Language, 25,* 555–584.

Genesee, F., Nicoladis, E., & Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language, 22,* 611–631.

Klee, T., Stokes, S. F., Wong, A. M.-Y., Fletcher, P., & Gavin, W. J. (2004). Utterance length and lexical diversity in Cantonese-speaking children with and without SLI. *Journal of Speech, Language and Hearing Research, 47,* 1396–1410.

Krashen, S. D., & Terrell, T. D. (1983). *The natural approach: Language acquisition in the classroom.* Oxford, England: Pergamon.

Lai, R. Y.-K. (2005, March). *Asymmetrical code-mixing in a case study of uneven development: Bootstrapping hypothesis or ivy hypothesis?* Paper presented at the Fifth International Symposium on Bilingualism, Barcelona, Spain.

Lakshmanan, U. (1995). Child second language acquisition of syntax. *Studies in Second Language Acquisition, 17,* 301–329.

Lanza, E. (2004). *Language mixing in infant bilingualism: A sociolinguistic perspective.* Oxford, England: Oxford University Press.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Matthews, S. & Yip, V. (1994). *Cantonese: A comprehensive grammar.* London: Routledge.

Matthews, S., & Yip, V. (2003). Relative clauses in early bilingual development: Transfer and universals. In A. G. Ramat (Ed.), *Typology and second language acquisition* (pp. 39–81). Berlin, Germany: de Gruyter.

Müller, N. (1998). Transfer in bilingual first language acquisition. *Bilingualism: Language and Cognition, 1,* 151–171.

Packard, J. (2000). *The morphology of Chinese: A linguistic and cognitive approach.* Cambridge, England: Cambridge University Press.

Paradis, J., & Genesee, F. (1996). Syntactic development in bilingual children: Autonomous or independent? *Studies in Second Language Acquisition, 18,* 1–25.

Petersen, J. (1988). Word-internal code-switching constraints in a bilingual child grammar. *Linguistics, 26,* 479–493.

Romaine, S. (1995). *Bilingualism* (2nd ed.). Oxford, England: Blackwell.

Saunders, G. (1988). *Bilingual children: From birth to teens.* Clevedon, England: Multilingual Matters.

Sgall, P. (1995). Prague school typology. In M. Shibatani & T. Bynon (Eds.), *Approaches to language typology* (pp. 49–84). Oxford, England: Clarendon.

Swain, M., & Wesche, M. (1975). Linguistic interaction: Case study of a bilingual child. *Language Sciences, 37,* 17–22.

Yip, V., & Matthews, S. (2000). Syntactic transfer in a Cantonese–English bilingual child. *Bilingualism: Language and Cognition, 3,* 193–207.

Yip, V., & Matthews, S. (2005). Dual input and learnability: Null objects in Cantonese–English bilingual children. In *Proceedings of the Fourth International Symposium on Bilingualism* (pp. 2421–2431). Tempe: Arizona State University.

Yiu, S.-M. E. (2005, March). *Asymmetrical language mixing in a Cantonese–English bilingual child.* Paper presented at the Fifth International Symposium on Bilingualism, Barcelona, Spain.