香港中文大學
The Chinese University of Hong Kong

兒童雙語研究中心
Childhood Bilingualism Research Centre

香港大學
THE UNIVERSITY OF HONG KONG

Code-mixing, complexity and translanguaging in bilingual children

Stephen Matthews

Emeritus Professor, University of Hong Kong

Co-director, Childhood Bilingualism Research Centre, Chinese University of Hong Kong

Dr. Jane Lai (Postdoctoral Fellow), CUHK

Prof. Maggie Mai, CUHK

Collaborators

Prof. Virginia Yip, CUHK

Research questions

- Are children's code-mixed utterances more complex than their monolingual utterances?
- If so, why and in what respects are they complex?
- What implications might this complexity have?
- How does it relate to translanguaging in Applied Linguistics?

Traditional/ Public Views of Code-mixing

"The use of more than one language in a single utterance" (Yip & Matthews, 2016)

Traditional / Public Views

A sign of confusion/ language interference:

Children failed to differentiate between the languages (e.g. Redlinger & Park, 1980; Volterra & Taeschner, 1978)

A "threat" to good English/ the purist view of "one language at a time":

Mixed code teaching: a common practice in Hong Kong schools, yet discouraged by the Education Bureau (Low & Lu, 2006)

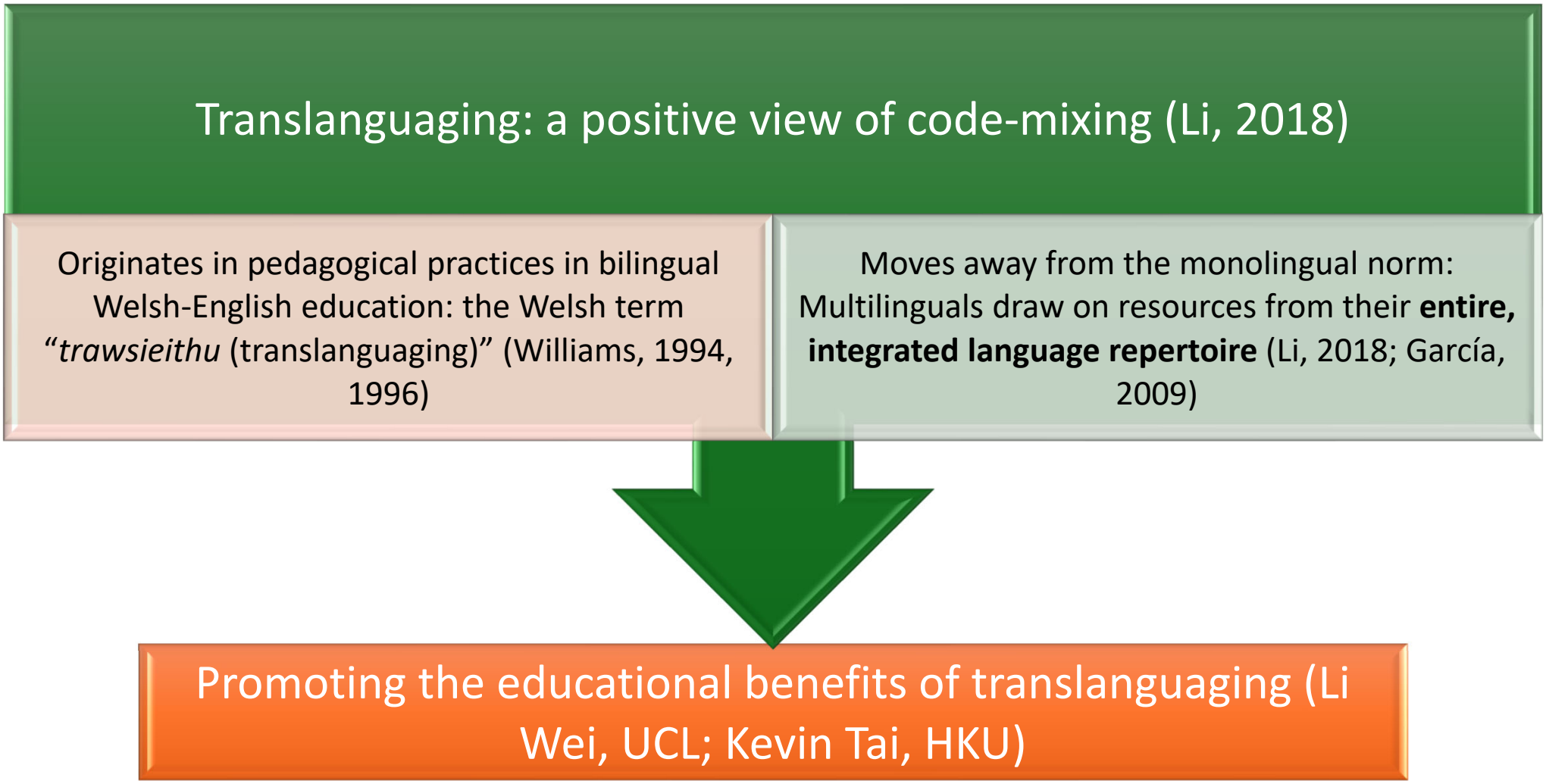
Public views of code-mixing

- Recommendation of Education Commission:

(iii) the use of mixed-code in schools should be reduced in favour of the clear and consistent use in each class of Chinese or English in respect of teaching, textbooks and examinations. (Education Commission, 1990: 99)

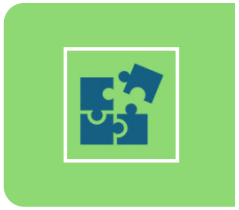
- Study commissioned by government failed to show expected result, not published (Antony Sweeting, p.c.)

Code-mixing as Translanguaging



Emerging evidence for a positive view of code-mixing

Linguistic Competence: code-mixing/switching is positively related to language competency, supporting the development of both dominant and weaker language (Yow et al. 2018; Schächinger Tenés et al., 2023)



Linguistic Creativity: bilingual Cantonese-English speaking children's code-mixing patterns go beyond the input and demonstrate creative construction (Yip & Matthews, 2016)

Cognitive Advantages: code-mixing/ switching is positively related to enhanced executive functioning skills in bilinguals (Hartanto & Yang, 2020)

Higher Complexity: mixed utterances are more complex than monolingual utterances- MLU was higher in mixed utterances than monolingual German or English utterances in three German-English bilingual children from 2;3 to 3;11 (Quick et al., 2018)

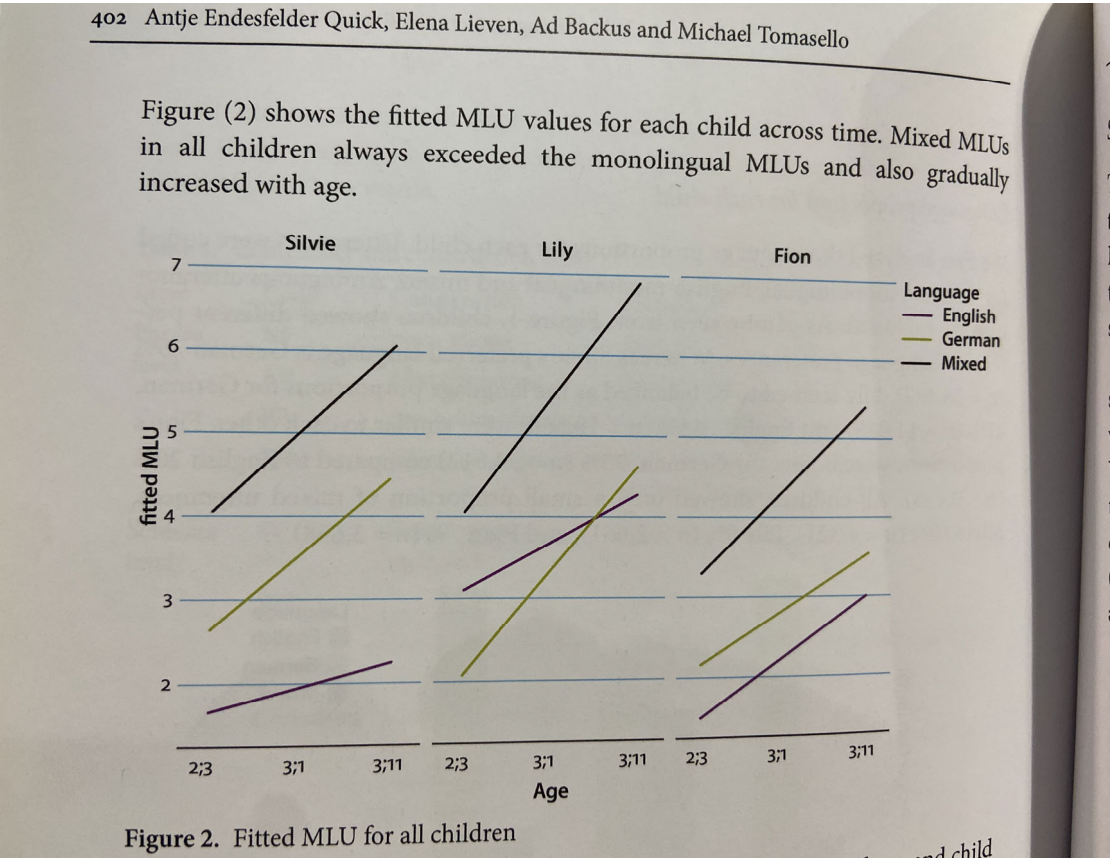
Mean Length of Utterance (MLU) in bilingual children's code-mixed vs. monolingual speech

- Yiu (2005): case study of code-mixing in bilingual child Kathryn (age 3;01-4;06)
- calculated Kathryn's MLU for Cantonese and English
- Discovered MLU for mixed utterances higher than either Cantonese or English
- Not included in thesis (!)
- Code-mixed utterances excluded for the purpose of calculating MLU differentials (as a measure of language dominance: Yip & Matthews 2006)
- Concern over comparability

Mean Length of Utterance (MLU) in bilingual children's code-mixed vs. monolingual speech

- Quick et al (2018) compared complexity of children's code-mixed and monolingual speech using **MLU, syntactic completeness and construction types**
- Code-mixed utterances showed **significantly higher MLU** than monolingual German or English utterances
- Methodological issue: code-mixed utterances must contain **at least 2 words** (or morphemes, if using MLUm) whereas monolingual utterances include single-word utterances
- Implication: the complexity advantage of code-mixed speech may be **over-estimated**

Mean Length of Utterance (MLU) in bilingual children's code-mixed vs. monolingual speech



402 Antje Endesfelder Quick, Elena Lieven, Ad Backus and Michael Tomasello

Figure (2) shows the fitted MLU values for each child across time. Mixed MLUs in all children always exceeded the monolingual MLUs and also gradually increased with age.

Figure 2. Fitted MLU for all children

- Code-mixed utterances of 3 children showed **significantly higher MLU** than monolingual German or English utterances

- But:
- MLU values for English begin below 2

-> Single-word utterances are included

- Mixed utterances cannot have MLU values below 2
- Comparison exaggerates the MLU advantage for mixed utterances

Minimal utterances

Monolingual:
one word (or morpheme)

搽
caa4
spread
‘put some (lipstick) on’

wear
‘put (clothes) on’

Mixed:
two words (or morphemes)

you 搽
you caa4
you spread
‘put some (lipstick) on’ (Alicia, 2;10;29)

wear 咗
‘(I) put (the clothes) on’

Current Study

- **Corpus study of nine bilingual Cantonese-English speaking children (Yip & Matthews, 2007):**
  - Following Quick et al., (2018), we compared the MLU for mixed versus monolingual utterances
  - **Methodological modifications:**

    i. For fair comparisons, only **multi-word utterances** were considered: compare mixed and monolingual utterances with >1 words

    ii. Mixed and monolingual utterances in both languages were extracted **regardless of the language elicitation context** (‘Cantonese’ vs ‘English’)

Data

- Hong Kong Bilingual Child Language Corpus in CHILDES: 9 Cantonese-English bilingual children aged between 1;03 – 4;06

	Age	Languages (Input Source)	Home Language Policy
Alicia	1;03;10 – 3;00;24	Cantonese (mother) English (father)	One-Parent One-Language
Charlotte	1;08;28 – 3;00;03	Cantonese (mother) English (father)	One-Parent One-Language
Darren	1;07;23 – 3;11;24	Cantonese (mother & father)	One-Parent Two-Languages
Janet	2;10;16– 3;11;11	Cantonese (mother) English (father)	One-Parent One-Language
Kasen	2;04;07 – 4;00;09	Cantonese (mother & father)	One-Parent Two-Languages
Kathryn	3;01;05 – 4;06;07	Cantonese (father) English (mother)	One-Parent One-Language
Llywelyn	2;00;12 – 3;04;17	Cantonese (mother) English (father)	One-Parent One-Language
Sophie	1;06;01 – 3;00;09	Cantonese (mother) English (father)	One-Parent One-Language
Timmy	2;00;26 – 3;06;25	Cantonese (mother) English (father)	One-Parent One-Language

Methodology

- Mixed and monolingual Cantonese and English utterances (n= 72,860) were extracted using CLAN commands.
  - Instances of code-mixing are labelled with a postcode @s in the transcripts:

    帶 佢 去 picnic@s 呀
    'take him for a picnic' (Timmy 3;5;14)

  - Proper names are excluded

Methodology

- **MLUw (in words)** was computed for mixed and monolingual **multi-word** utterances.
- Compare MLUw for:
  - **Monolingual Cantonese multi-word utterances**
  - **Monolingual English multi-word utterances**
  - **Mixed utterances**

Results: mean MLUw in each language across all children

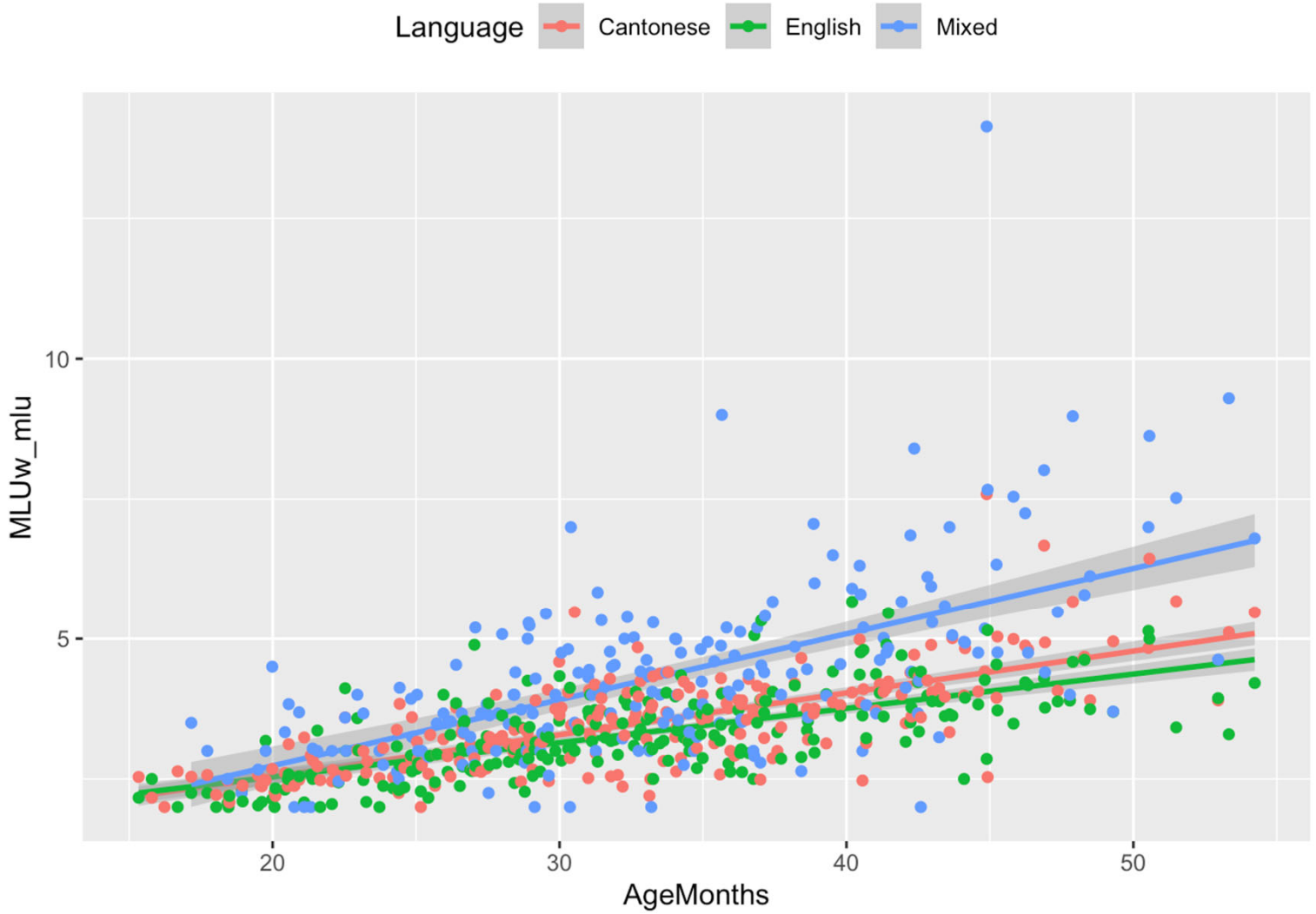
	Mean MLUw of Cantonese multi-word utterances	Mean MLUw of English multi-word utterances	Mean MLUw of Mixed utterances
All children (n=9)	3.48	3.33	4.38

Data Analysis

1) Linear mixed effects model fitted for all children, with MLUw as the dependent measure, and language (Can/Eng/Mixed), age and their interactions as predictors.

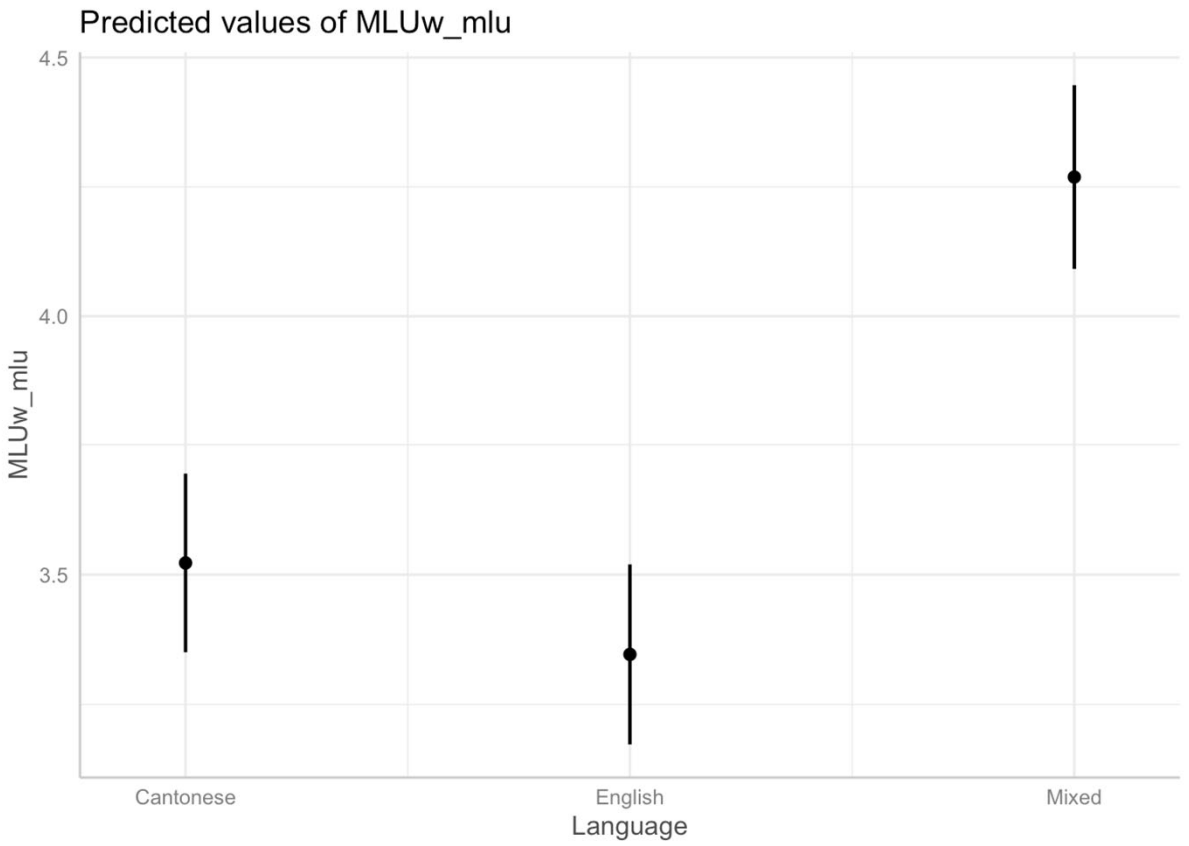
2) Linear regression model fitted for each child, with MLUw as the dependent measure, and language (Can/Eng/Mixed), age and their interactions as predictors.

MLUw by Age in all children



Findings across all children (n=9)

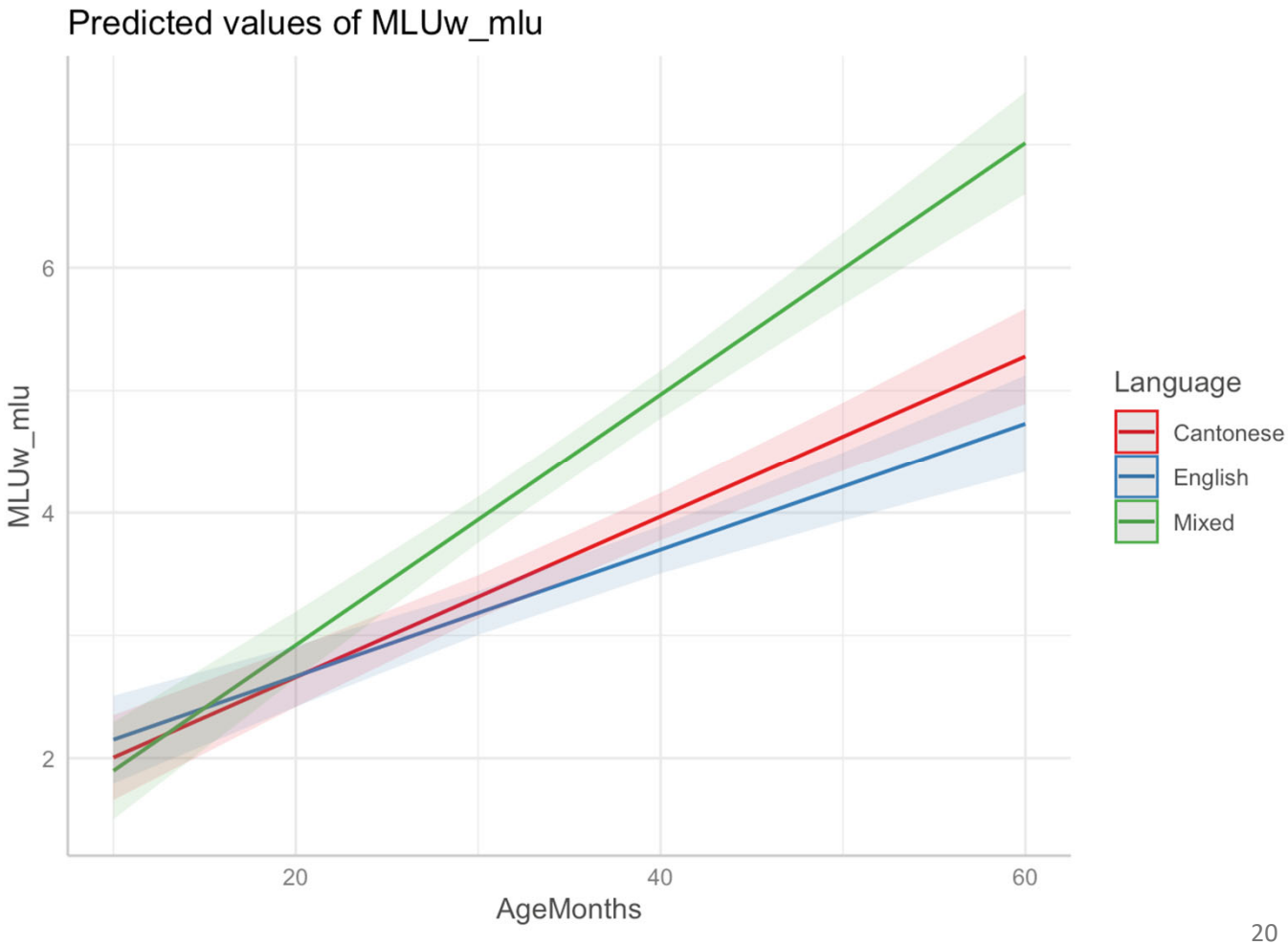
- Sig. main effect of Language***
- *Posthoc emmeans analyses:*
  - Cantonese > English*
  - Cantonese < Mixed***
  - English < Mixed***



Findings across all children (n=9)

- Sig. interaction between Language and Age*** :

The older the child, the greater the MLU advantage in **Mixed over Monolingua**l Cantonese/ English utterances.



Data Analysis

1) Linear mixed effects model fitted for all children, with MLUw as the dependent measure, and language (Can/Eng/Mixed), age and their interactions as predictors.

2) Linear regression model fitted for each child, with MLUw as the dependent measure, and language (Can/Eng/Mixed), age and their interactions as predictors.

Child	Linear Regression Model	Post-hoc emmeans analyses
Alicia	Sig. main effect of Language* Sig. main effect of Age***	Cantonese > English** Cantonese < Mixed* English < Mixed***
Charlotte	Sig. main effect of Language*** Sig. main effect of Age***	Cantonese < English* Cantonese < Mixed*** English < Mixed*
Darren	Sig. main effect of Language*** Sig. main effect of Age***	Cantonese = English Cantonese < Mixed*** English < Mixed**
Janet	Sig. main effect of Age**	N/A
Kasen	Sig. main effect of Language*** Sig. main effect of Age***	Cantonese = English Cantonese < Mixed*** English < Mixed***
Kathryn	Sig. main effect of Language***	Cantonese = English Cantonese < Mixed*** English < Mixed***
Llywelyn	Sig. main effect of Language* Sig. main effect of Age***	Cantonese = English Cantonese < Mixed* English < Mixed**
Sophie	Sig. main effect of Language** Sig. main effect of Age***	Cantonese = English Cantonese < Mixed** English < Mixed***
Timmy	Sig. main effect of Language*** Sig. main effect of Age*** Sig. interaction between Language and Age**	Cantonese = English Cantonese < Mixed*** English < Mixed***

Summary of findings

- At the group level, **mixed utterances were significantly longer** than monolingual Cantonese and English multi-word utterances.
- 8 out of 9 children demonstrated **significantly higher MLU** in mixed versus monolingual Cantonese/English utterances.
- As a group: the older the child, the greater the MLU advantage in mixed over monolingual Cantonese/English utterances.

Discussion

- Why are mixed utterances longer than monolingual ones?
- In what respects are mixed utterances more complex?
- Do the findings have educational implications?

Discussion

1) Reiteration of meanings across Cantonese and English within the same utterance:

banana 香蕉 (Alicia; 1;6;28)

Banana hoeng1ziu1

I want 要 like this (Alicia; 1;10;29)

I want jiu3 like this

攞 鐘 clock 呀 (Charlotte, 2;0;25)

lo2 zung1 clock aa4

'get a clock'

- Reiteration (Gumperz 1982) not only leads to longer utterances, but also highlights children's pragmatic ability to emphasize or clarify their intended meanings

Discussion: prevalent types of mixed utterances

2) Insertion of sentence final particles (SFPs):

house aa4 (Charlotte; 1;8;28)

I go laa3 (Charlotte, 2;5;19)

I think we need to see the tv aa4 (Janet; 3;8;25)

you eat laa1 (Kasen 2;9;1)

bees lai4 gaa3 (Alicia; 2;4;24)

pumpkin lo1 (Timmy 2;10;7)

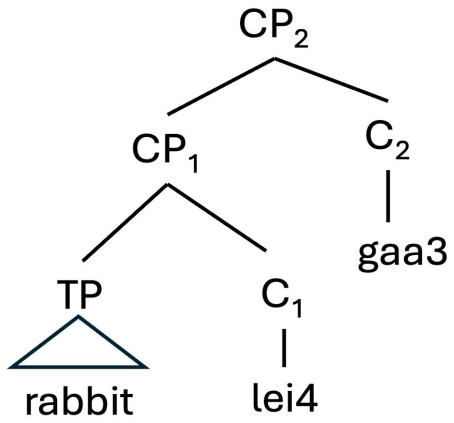
- (why) do SFPs add complexity?

Sentence particles: structural complexity

- Particles are assumed to head a functional projection such as CP (Sybesma & Li 2007, Simpson 2014) or FP (Force Phrase, Lam 2014) with at least two positions

Rabbit lei4gaa3

‘ ít’s a rabbit’ (Alicia 2;7;28)



Sentence particles: pragmatic complexity

- A reflection of children's pragmatic competence and discourse strategies to express nuanced pragmatic meanings

Rabbit lei4gaa3 Alicia 2;7;28)

'It's a rabbit' (affirmative/explanatory)

I go laa3 (Charlotte, 2;5;19)

'I'm going now,' 'I'm off!' (current relevance)

Watch tv lo1 (Kasen, 3;0;3)

'(I'm) watching TV, aren't I?' (evidential/impatient)

Discussion

3) Verb-particle constructions:

我　　想　　turn off　呀

Ngo5 soeng2 turn off aa3

I　　want　　turn off SFP

‘I want to turn (it) off’ (Charlotte, 2;4;20)

我　　幫　　你　put it

Ngo5 bong1 nei5 put it

I　help　you　put　it

‘I’ll help you put it in place’ (Charlotte; 2;0;25)

跟住 呢，佢 throw away

Gan1zyu6 le1, keoi5 throw away

Following SFP, he throws away

‘Then he throws it away’ (Kasen 3;5;5)

依隻　slide　唔　倒　　down 呀

Ji1zek slide m4 dou2 down aa3.

This CL slide not succeed down SFP

‘With this (shoe) one can’t slide down properly’

(Timmy; 2;11;20)

Discussion

4) complementation: embedded clause

I want 街街
I want gaai1gaai1
I want street
‘I want to go out’ (Alicia; 2;10;29)

I want 睇 呢 隻
I want tai2 ni1 zek3
I want see this CL
‘I want to see this’ (Alicia; 2;10;29)

我 想 open (Darren 2;7;8)
‘I want to open (it)’

我 幫 你 put it (Charlotte 2;0;25)
‘I want you put it (in)’

我 唔 想 be caught 呀 嘛
‘I don’t want to be caught’ (Kathryn 4;2;17)

Code-mixing as translanguaging

The **increasingly higher MLU in mixed utterances** suggests that as these children became more competent with both languages, they **creatively combine linguistic resources to express complex content**.

Children's code-mixing is reflective of their **communicative and social competence** in employing all their linguistic resources to convey meanings and to facilitate learning and communication (Genesee, 2006).

Consistent with cross-linguistic findings (e.g. Yow et al., 2018, Quick et al., 2018); and the translanguaging perspective (Li, 2018)

Beyond bilingualism:
Translanguaging in three languages

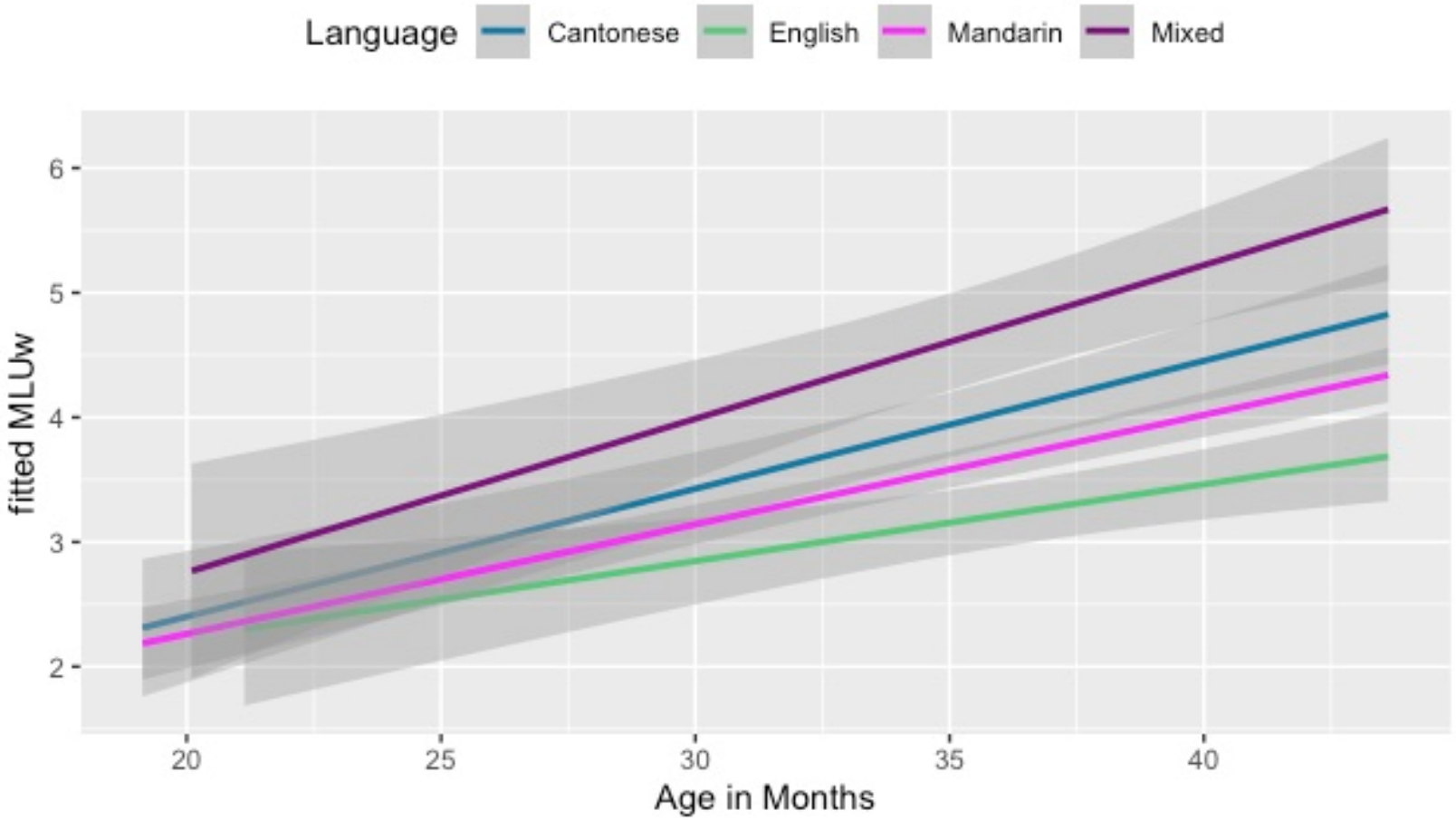
Trilingual corpora (Mai & Yip 2017, 2022)

	Winston	Leo	Louis
Age	1;07-3;07	1;06-2;11	2;0-2;11
Languages (Input Source)	Cantonese (mother, grandparents) English (schooling) Mandarin (father, grandparents)	Cantonese (mother) English (mother, helper, teachers) Mandarin (father, paternal grandmother)	Cantonese (mother) English (mother, helper) Mandarin (father, paternal grandmother)
Ambient Language (location)	English (United States)	Cantonese (Hong Kong)	Cantonese (Hong Kong)
Home Language Policy	One-Parent One-Language (though not strictly); Grandparents: 70% Cantonese; 30% Mandarin	One-Parent One-Language before 1;01; One-day One-Language after 1;01	One-Parent One-Language before 1;01; One-day One-Language after 1;01

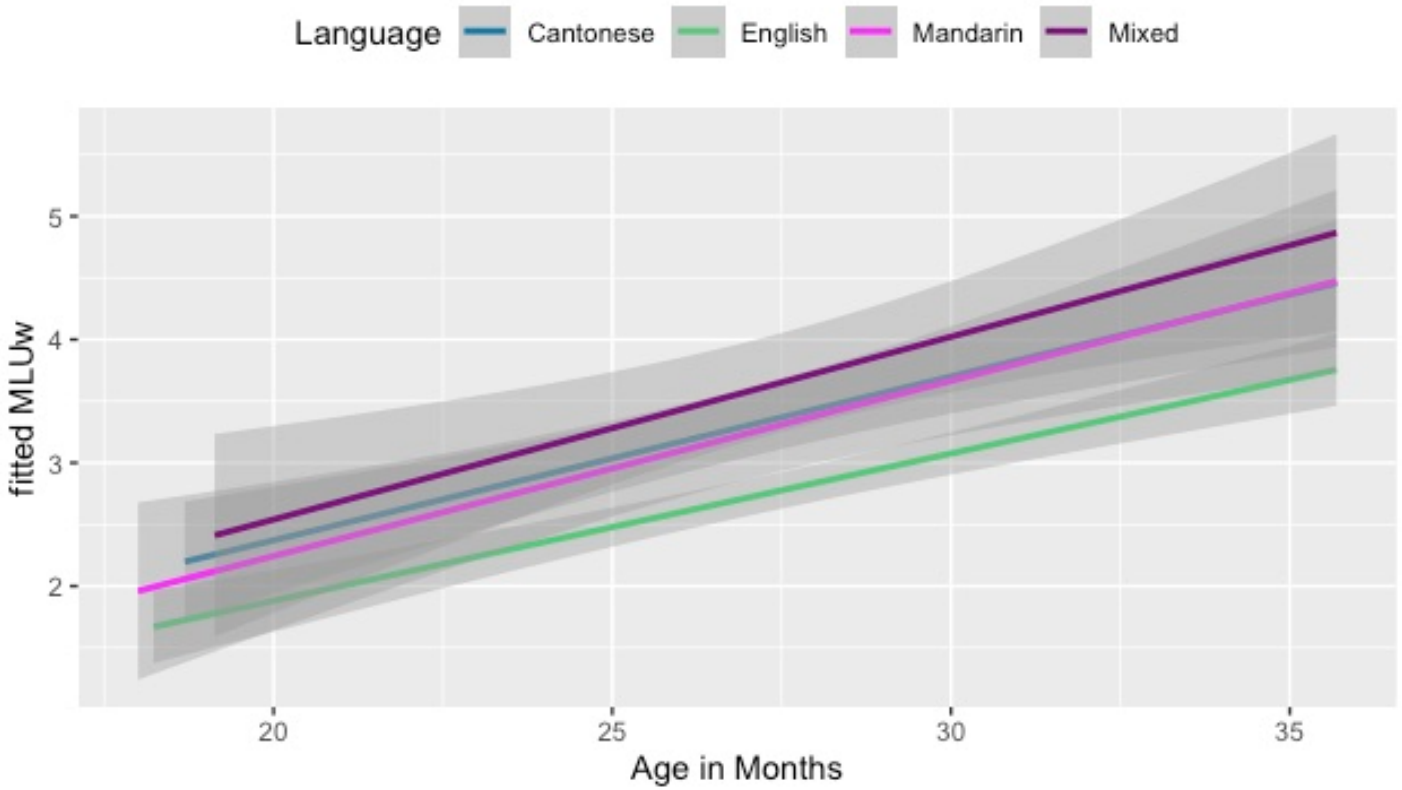
Mean MLUw of each utterance type in the three trilingual children

	Cantonese multi-word utterances	English multi-word utterances	Mandarin multi-word utterances	Mixed utterances
Winston	3.97	3.28	3.74	4.95
Leo	3.25	2.75	3.16	3.59
Louis	2.81	2.90	3.00	4.18

Fitted MLUw of mixed and monolingual utterances for Winston



Fitted MLUw of mixed and monolingual utterances for Leo



Beyond bilingualism: Translanguaging in three languages

中　文　可　能　講　唔　到　kitchen

Zhong1wen2 ho2lang4 gong2 m4 dou2 kitchen

‘Maybe we can’t say kitchen in Chinese’ (Winston 3;5;10)

- Lexical gaps recognised/hypothesized by child (Liang 2024)
- Metalinguistic commentary: from age 2 in bilinguals, after age 4 in monolinguals

Gong2 Ciu4zau1 Wa2, is “puat lok k'u". Fall down.

‘Talking Chiu Chow, it’s “puat lok k’u” (meaning) “fall down”’ (Timmy 2;02;03)

Summary

- Code-mixed utterances are consistently more complex than in Cantonese-English bilingual and Cantonese-English-Mandarin trilingual children
- The advantage of mixed utterance grows during the preschool years
- Consistent with translanguaging perspective: bilinguals pool resources to express more complex content
- Implications: advice for parents, teacher training, educational policy

References

- Bernardini, P., & Schlyter, S. (2004). Growing syntactic structure and code-mixing in the weaker language: The ivy hypothesis. *Bilingualism: Language and* Cognition, 7, 49–69. https://doi.org/10.1017/S1366728904001270
- Deuchar, M. & Quay, S. (2000). *Bilingual acquisition: Theoretical implications of a case study*. Oxford: Oxford University Press.
- Education Commission (1990). Report no. 4: The Curriculum and Behavioural Problems in Schools. Hong Kong Education Commission.
- García, O. (2009). *Bilingual education in the 21st century: A global perspective.* Oxford: Wiley-Blackwell.
- Genesee, F. (2006). Bilingual first language acquisition in perspective. In P. McCardle & E. Hoff (Ed.), *Childhood Bilingualism: Research on Infancy through School Age* (pp. 45-67). Bristol, Blue Ridge Summit: Multilingual Matters. https://doi.org/10.21832/9781853598715-005
- Hartanto, A., & Yang, H. (2020). The role of bilingual interactional contexts in predicting interindividual variability in executive functions: A latent variable analysis. *Journal of Experimental Psychology: General, 149*(4), 609–633.
- Hoffmann, C. (2000). Bilingual and trilingual competence: Problems of description and differentiation. *Estudios de Sociolingüística, 1*, 83-92.
- Lam, Z. (2014). A complex ForceP for speaker- and addressee-oriented discourse particles in Cantonese. Studies in Chinese Linguistics 35, 61-80.
- Lee, J. H. N., Lai, R. Y. K., Matthews, S., & Yip, V. (2023). Prosodic interaction in Cantonese-English bilingual children's speech production. *Linguistic Approaches to Bilingualism.* https://doi.org/10.1075/lab.21049.lee
- Leopold, W.F. (1978). A child's learning two languages. In E. Hatch (ed.), *Second language acquisition: A book of readings.* Rowley, MA: Newbury House, 23-32.
- Li, W. (2018). Translanguaging as a practical theory of language*. Applied Linguistics, 39*(1), 9-30. https://doi.org/10.1093/applin/amx039

References

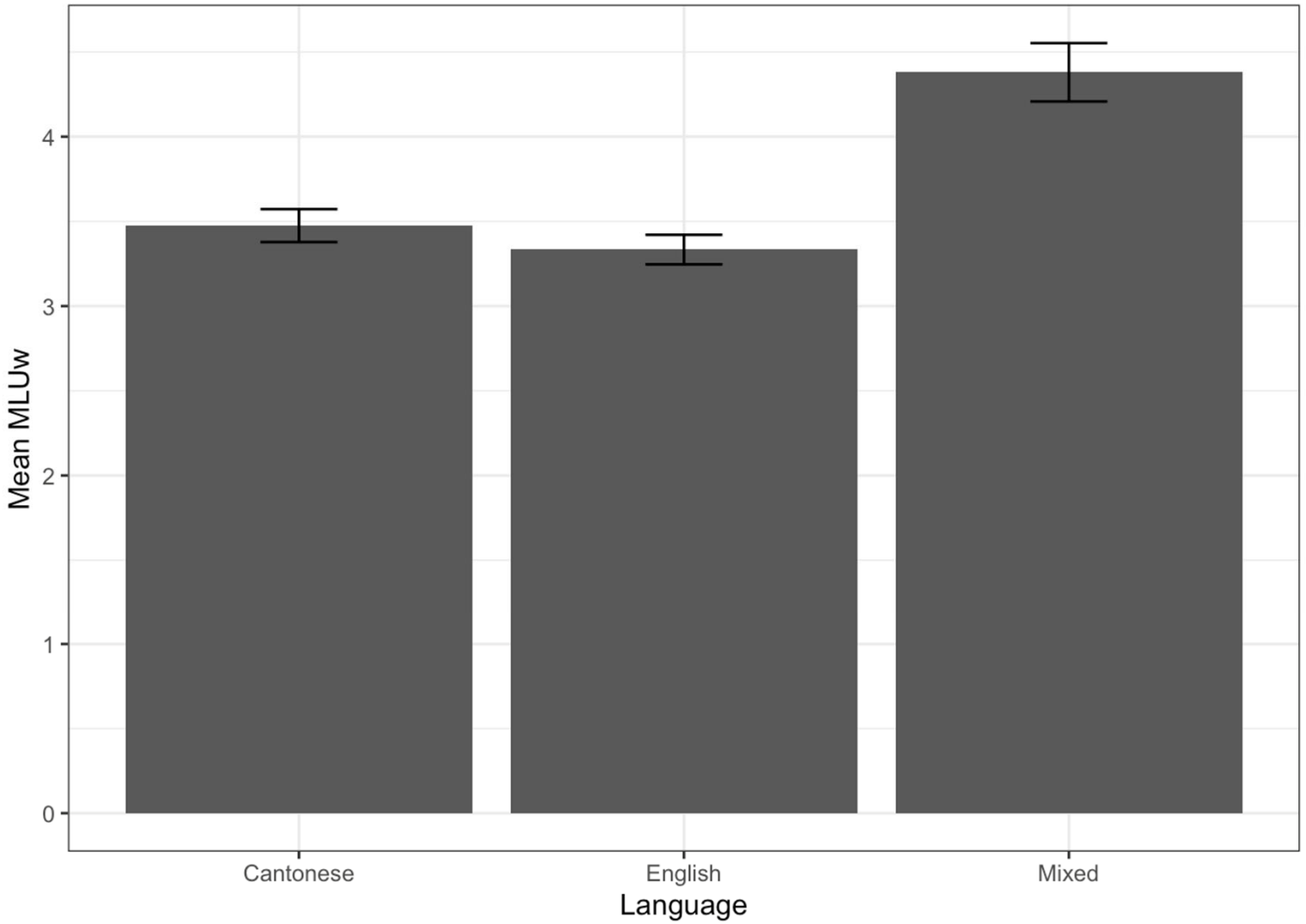
- Liang, M. W. (2024). Translanguaging in trilingual children: The lexical gap hypothesis. *Unpublished MA Thesis*, The Chinese University of Hong Kong.
- Low, W. W. M., & Lu, D. (2006). Persistent use of mixed code: An exploration of its functions in Hong Kong schools. *International Journal of Bilingual Education and Bilingualism, 9*(2), 181-204. https://doi.org/10.1080/13670050608668640
- Mai, Z. & Yip, V. (2017). Acquiring Chinese as a Heritage Language in English-speaking Countries and the Child Heritage Chinese Corpus. *Paper presented at the International Conference on Bilingualism: Language and Heritage,* Dec 18, Chinese University of Hong Kong.
- Mai, Z., & Yip, V. (2022). Caretaker input and trilingual development of Mandarin, Cantonese and English in early childhood (1;6-2;11). *International Journal of Bilingual Education and Bilingualism, 25*(9), 3389-3403. https://doi.org/10.1080/13670050.2022.2060037.
- Quick, A. E., Lieven, E., Backus, A., & Tomasello, M. (2018). Constructively combining languages: The use of code-mixing in German-English bilingual child language acquisition. *Linguistic Approaches to Bilingualism, 8*(3), 393-409. https://doi.org/10.1075/lab.17008.qui
- Quick, A. E., & Hartmann, S. (2021). The building blocks of child bilingual code-mixing: A cross-corpus traceback approach*. Frontiers in Psychology, 12*, 682838. https://doi.org/10.3389/fpsyg.2021.68283
- Redlinger, W. E., & Park, T. Z. (1980). Language mixing in young bilinguals*. Journal of Child Language, 7*(2), 337-352. https://doi.org/10.1017/S030500090000266X
- Schächinger Tenés, L. T., Weiner-Bühler, J. C., Volpin, L., Grob, A., Skoruppa, K., & Segerer, R. K. (2023). Language proficiency predictors of code-switching behavior in dual-language-learning children. *Bilingualism (Cambridge, England)*, *26*(5), 942–958. https://doi.org/10.1017/S1366728923000081

References

- Sybesma, R. & Li, B. (2007). The dissection and structural mapping of Cantonese sentence final particles. *Lingua* 117, 1739-1783.
- Volterra, V., & Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of Child Language, 5,* 311–326. https://doi.org/10.1017/S0305000900007492
- Williams, C. (1994). Arfarniad o Ddulliau Dysgu ac Addysgu yng Nghyddestun Ad-dysg Uwchradd Ddwyieithog [An evaluation of teaching and learning methods in the context of bilingual secondary education]. *Unpublished Doctoral Thesis,* University of Wales, Bangor.
- Williams, C. (1996). Secondary Education: Teaching in the bilingual situation. In C. Williams, G. Lewis & C. Baker (Eds.), *The language policy: Taking stock* (pp. 39-78). Llangefni, UK: CAI.
- Yip, V., & Matthews, S. (2006). Assessing language dominance in bilingual acquisition: A case for mean length utterance differentials*. Language Assessment Quarterly: An International Journal, 3*(2), 97-116. https://doi.org/10.1207/s15434311laq0302_2
- Yip, V., & Matthews, S. (2007). *The bilingual child: Early development and language contact.* Cambridge: Cambridge University Press.
- Yip, V., & Matthews, S. (2016). Code-mixing and mixed verbs in Cantonese-English bilingual children: Input and innovation. *Languages, 1*(1), 4. https://doi.org/10.3390/languages1010004
- Yow, W. Q., Tan, J. S. H., & Flynn, S. (2018). Code-switching as a marker of linguistic competence in bilingual children. *Bilingualism: Language and Cognition, 21* (5), 1075-1090. https://doi.org/10.1017/S1366728917000335

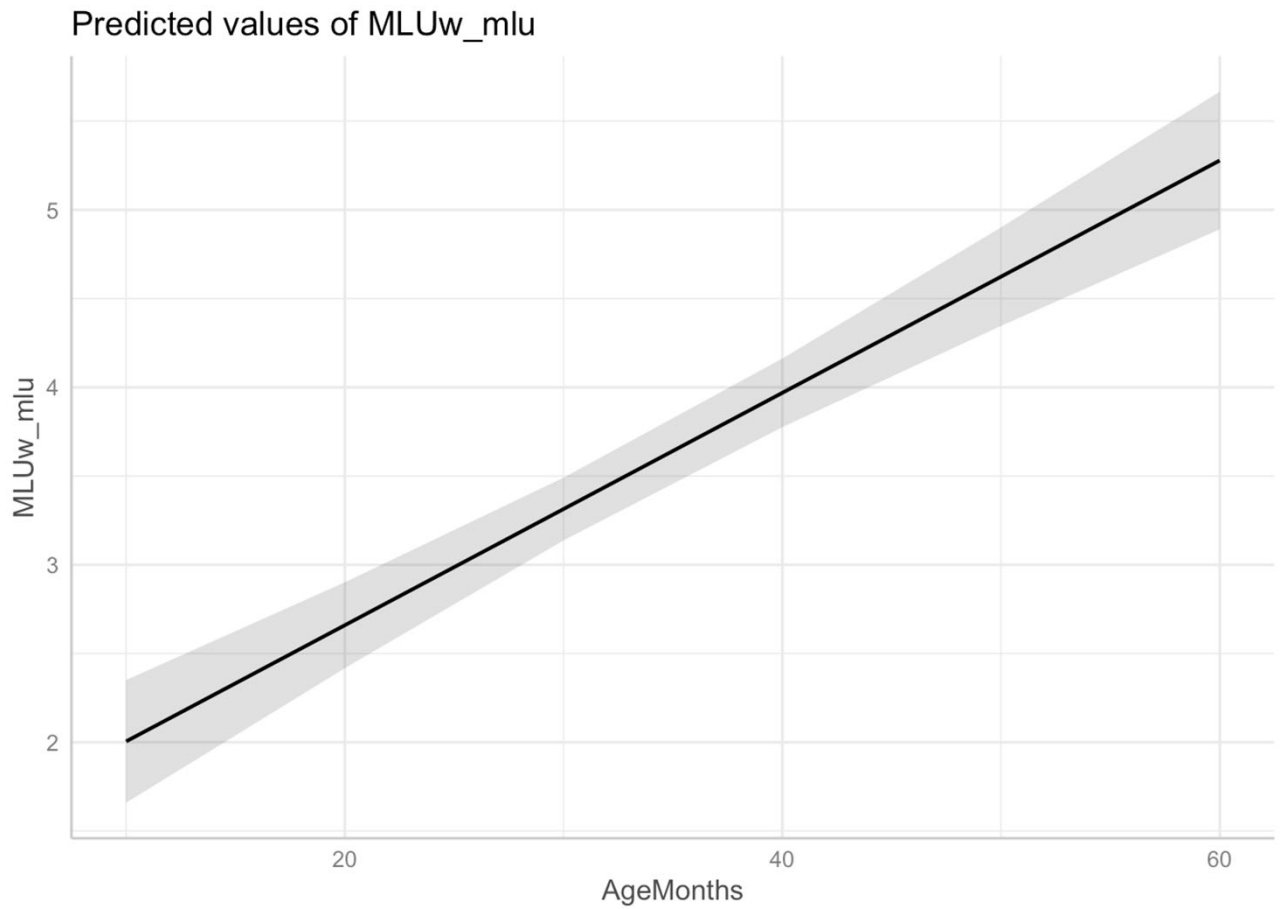
Supplementary Slides

Mean MLUw in each language across all children



Findings across all children (n=9)

- Sig. main effect of Age***



	Total no. of multi-word utterances	Cantonese multi-word utterances	English multi-word utterances	Mixed utterances
Alicia	n=6902	72.79% (n=5024)	24.01% (n=1657)	3.20% (n=221)
Charlotte	n=4565	27.69% (n=1264)	67.51% (n=3082)	4.80% (n=219)
Darren	n=6117	46.71% (n=2857)	50.30% (n=3077)	2.99% (n=183)
Janet	n=8024	69.80% (n=5601)	25.65% (n=2058)	4.55% (n=365)
Kasen	n=9173	44.51% (n=4083)	45.15% (n=4142)	10.34% (n=948)
Kathryn	n=6954	56.80% (n=3950)	40.18% (n=2794)	3.02% (n=210)
Llywelyn	n=5510	54.45% (n=3000)	43.36% (n=2389)	2.20% (n=121)
Sophie	n=12408	77.09% (n=9565)	22.07% (n=2739)	0.84% (n=104)
Timmy	n=13638	63.66% (n=8682)	33.61% (n=4584)	2.73% (n=372)
TOTAL	**n=73291**	**60.07% (n=44026)**	**36.19% (n=26522)**	**3.74% (n=2743)**

MLU in words (MLUw) versus morphemes (MLUm)

Consistent with our team's previous work (e.g. Yip & Matthews, 2000; 2006) and other cross-linguistic studies such as Swedish-French/ Swedish-Italian (Bernardini & Schlyter, 2004)

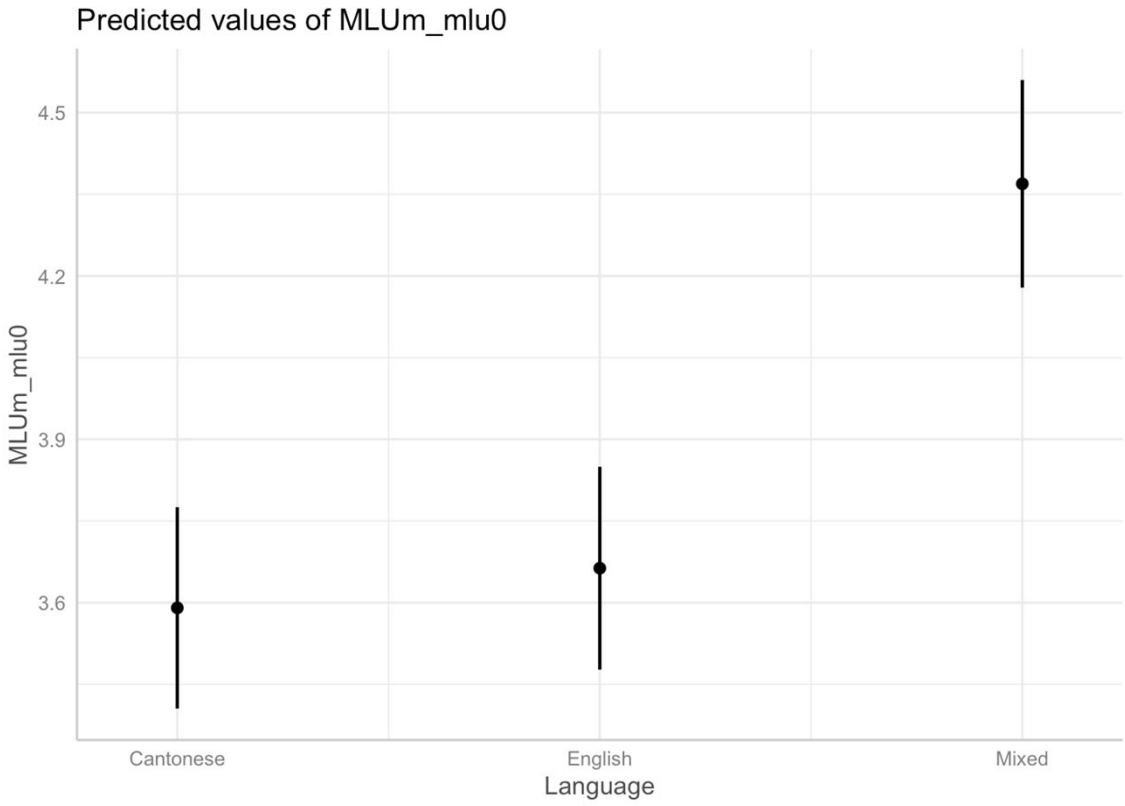
Typology of Cantonese as a (largely) isolating language, and child English could also be treated as **predominantly isolating**, given the lack of inflectional morphology produced in child speech (Yip & Matthews, 2000)

MLUw has been suggested as a **reliable measure** of Cantonese-speaking children's **morphosyntactic competence** (Klee et al., 2005)

Current study: MLUw and MLUm analyses yielded largely similar results for mixed versus monolingual multi-word utterances

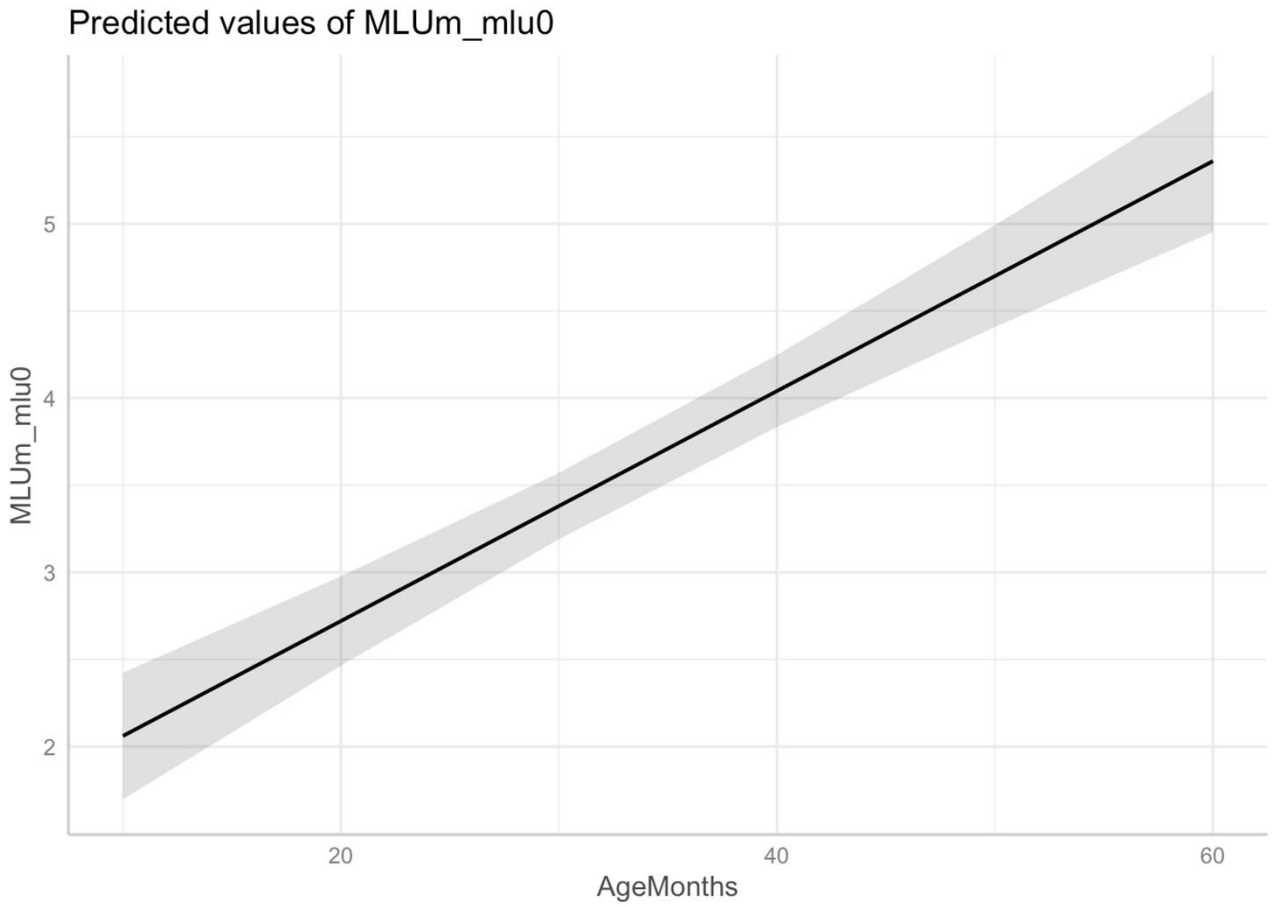
Analyses based on MLUm: Findings across all children (n=9)

- Sig. main effect of Language***
- *Posthoc emmeans analyses:*
  - Cantonese = English
  - Cantonese < Mixed***
  - English < Mixed***



Analyses based on MLUm: Findings across all children (n=9)

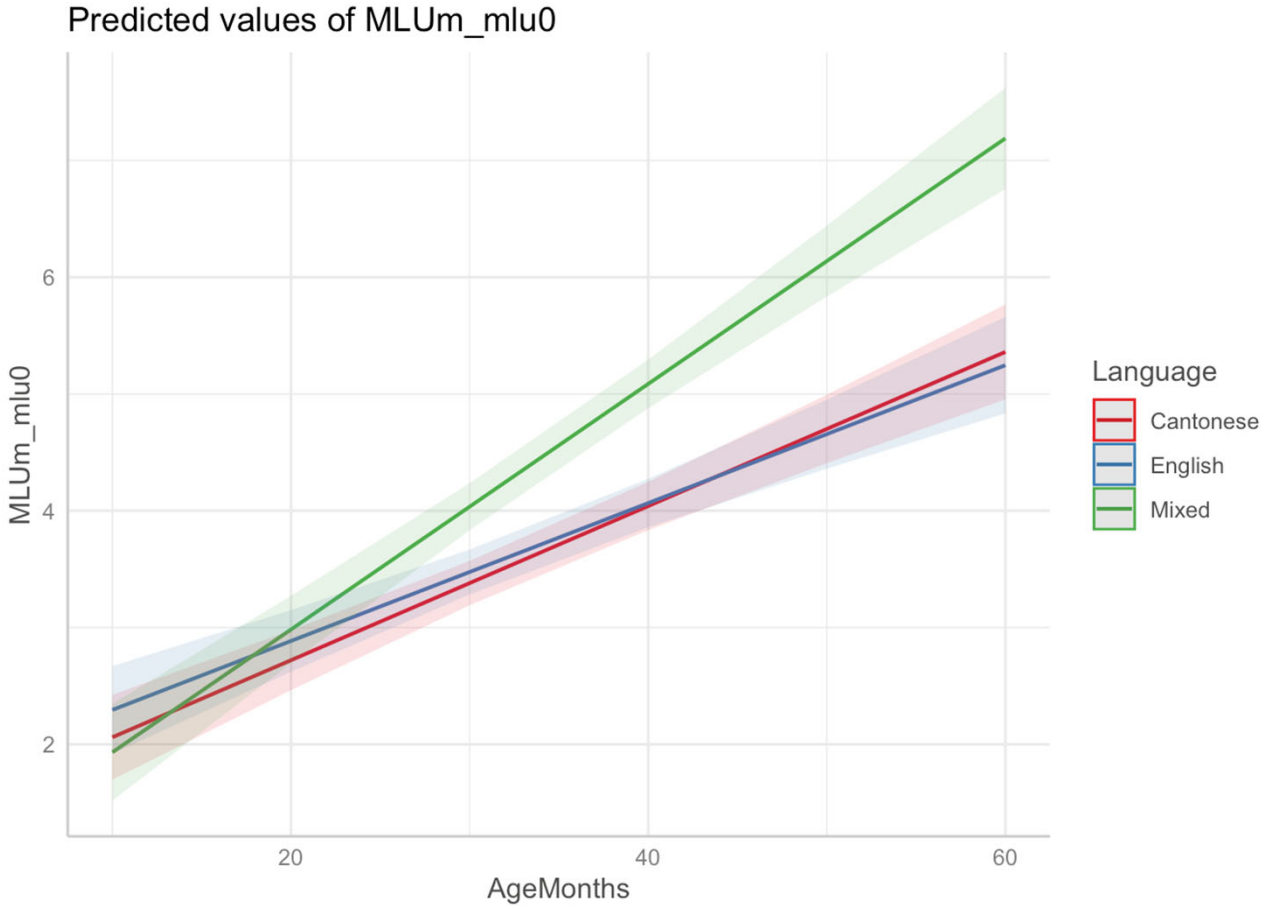
- Sig. main effect of Age***



Analyses based on MLUm: Findings across all children (n=9)

- Sig. interaction between Language and Age*** :

The older the child, the greater the MLU advantage in **Mixed over Monolingua**l Cantonese/ English utterances.



Child	Linear Regression Model	Post-hoc emmeans analyses
Alicia	Sig. main effect of Language* Sig. main effect of Age***	Cantonese > English* Cantonese < Mixed** English < Mixed***
Charlotte	Sig. main effect of Language*** Sig. main effect of Age***	Cantonese < English*** Cantonese < Mixed*** English = Mixed
Darren	Sig. main effect of Language*** Sig. main effect of Age***	Cantonese = English Cantonese < Mixed*** English = Mixed
Janet	Sig. main effect of Age**	N/A
Kasen	Sig. main effect of Language** Sig. main effect of Age**	Cantonese = English Cantonese < Mixed** English < Mixed*
Kathryn	Sig. main effect of Language***	Cantonese = English Cantonese < Mixed*** English < Mixed***
Llywelyn	Sig. main effect of Language** Sig. main effect of Age***	Cantonese = English Cantonese < Mixed* English = Mixed
Sophie	Sig. main effect of Language** Sig. main effect of Age***	Cantonese = English Cantonese < Mixed** English < Mixed**
Timmy	Sig. main effect of Language*** Sig. main effect of Age*** Sig. interaction between Language and Age**	Cantonese < English* Cantonese < Mixed*** English < Mixed**