Word comprehension and multilingualism among toddlers: A study using touch screens in daycares

Laia Fibla and **Charlotte Maniel** and **Alejandrina Cristia**
Laboratoire de Sciences Cognitives et Psycholinguistique (ENS, EHESS, CNRS)
Département d'Etudes Cognitives, Ecole Normale Supérieure, PSL Research University
29, Rue d'Ulm
Paris, 75005, FRANCE
{laia.fibla.reixachs|alecristia}@gmail.com

Abstract

Most previous research on young infants' spoken word comprehension has focused on monolinguals. These results may not generalize to non-monolingual populations because lexical processing may be more intricate for infants exposed to more than one language. Do toddlers learning multiple languages recognize words similarly to their monolingual peers? Answering this question will require extensive efforts, to which we contribute word comprehension data collected through a procedure aiming to be both precise and ecological. French-learning toddlers (N = 38; age range 1;11-3;4) were tested in their daycare, using a French-spoken prompt-to-picture matching task implemented on a child-friendly touch screen. Our results document some differences in accuracy, but not response time or number of trials completed, among toddlers differing in the number of languages they routinely hear. Additionally, these data suggest that it is feasible to collect good quality data from multiple children tested at once in daycares, opening the path to larger-scale studies. Future research could disentangle the many factors that are often empirically confounded with monolingual versus bilingual/multilingual status.

1 Introduction

Studying language acquisition among children exposed to multiple languages is interesting for both scientific and societal reasons. As to the former, studies on bilingual and multilingual infant language development could provide answers regarding the cognitive organization of language, and further our understanding of the system underlying early language acquisition. To take one example, there is an ongoing discussion regarding the role of the quantity of input directly addressed to the child in lexical development (Shneidman and Goldin-Meadow, 2012). Children exposed to more than one language will on average hear fewer words in each than a monolingual peer - that is, provided all else is equal among the two.

However, all else is frequently *not* equal when comparing monolingual children and those routinely exposed to more than one language. To begin with, while many bi- and multi-lingual children are exposed simultaneously to more than one language from birth, for others the acquisition of their additional language(s) occurs later on, including once the first language is already well established (de Houwer, 2009). Particularly in the latter cases, hearing multiple languages is correlated with immigrant status, itself correlated with lower education and income levels – which is itself correlated with lower levels of parental speech to the child (Hoff, 2013). As a result, studying acquisition in such populations is both complicated, and particularly relevant if one would hope to strive for equality in educational opportunities (Hoff, 2013). These arguments also underline the profound societal relevance of studies on bilingual acquisition.

Laia Fibla, Charlotte Maniel and Alejandrina Cristia 2016. Word comprehension and multilingualism among toddlers: A study using touch screens in daycares. *Proceedings of the joint workshop on NLP for Computer Assisted Language Learning and NLP for Language Acquisition at SLTC 2016*. Linköping Electronic Conference Proceedings 130: 18–23.

Setting these confounds aside for the moment, there is mounting evidence that monolinguals and non-monolinguals differ in several ways when lexical development is considered. Specifically, vocabulary sizes in monolingual preschoolers are comparable to their bilingual peers when both languages are combined together, and translation equivalents are considered only once (Pearson et al., 1997). However, when lexica are evaluated separately, bilingual children typically score lower than monolingual on vocabulary tests targeting their common language (Bialystok et al., 2010). These differences in vocabulary size could be due to a number of reasons, but by and large it appears that they emerge due to differences in input quantity: as noted previously, when exposed to several languages, bilingual children receive less input from each language than their monolingual peers (Place and Hoff, 2011). Such differences in vocabulary size have been documented with measures ranging from free production to standardized tests of lexical comprehension administered by an experimenter.

Further research suggests that bilinguals as a population vary on speed of lexical comprehension in ways that may relate to their vocabulary size. Marchman and colleagues used a looking-while-listening task, where two visual referents are presented on the screen and a spoken prompt asks for one of them (Marchman et al., 2010). Bilingual toddlers were tested in both their languages, and their parents were asked to complete a vocabulary checklist for each language. The authors document a significant correlation between speed of word recognition and vocabulary size within languages but not across them; i.e., individual variation in parental report of English vocabulary size predicted speed of recognition for English, but not Spanish, items, and *vice versa*. Since bilinguals have a smaller vocabulary than monolinguals when languages are considered separately, it appears likely that they will also be slower to recognize words. Although this contrast has not been carried out directly, comparison with other work published by the same group confirms the prediction. Whereas bilingual 30-month-olds averaged 860-870ms response times in either language, Spanish-learning monolinguals averaged 900ms when tested much earlier, at 24 months (Hurtado et al., 2008); and English-learning monolinguals displayed these response times even younger, at 18 months (Fernald et al., 1998).

In sum, previous research on lexical processing suggests that bilinguals are at a disadvantage compared to monolinguals. However, one recent study using a laboratory-based touch-screen test found that Canadian 2-year-olds' performance (accuracy and response time) in the common language did not differ as a function of bilingual status (Poulin-Dubois et al., 2013). It is possible that this touch-screen task is a more accurate index of children's competence, since it has been separately found to be a better predictor of later vocabulary than e.g. parental report (Friend and Keplinger, 2008). If so, other previous work may have underestimated bilinguals' lexical abilities. An alternative explanation holds that bilinguals and monolinguals in this sample may have been better matched on confounded variables, such as parental socio-economic status, than the American samples who are more commonly studied. A final alternative is that the difference does exist but is small enough to sometimes yield false negatives. Indeed, Poulin-Dubois and colleagues briefly point out that the difference is present numerically, but may not have surfaced due to limitations in statistical power. Teasing these three alternatives apart requires gathering additional data with a similar touch-screen test.

Our study sought to contribute to the general line of research documenting the relationship between word comprehension and exposure to multiple languages. For that purpose we tested monolingual, bilingual, and multilingual French toddlers *in their common language*. Our goal was not to assess global lexicon, but only to compare them on the one language they all shared. We were inspired by Friend et al. (2008)'s Computerized Comprehension Task (CCT) when developing a two-alternative prompt-to picture matching test on an iPad®, and thus followed their lead on most methodological choices (Friend and Keplinger, 2008). Most saliently, we had numerous trials sampling from three word categories (nouns, verbs, and adjectives) and three levels of difficulties (described in detail in the Methods) so that our measure was more representative of word processing in general. Unlike Friend and colleagues, we used portable technology so as to bring the test to the childrens natural environment, the daycare, a setting which may also have leveling properties by exposing all children (monolinguals and non-monolinguals) to similar experiences. By testing in three daycares located in the $13th$ neighborhood in Paris, France, we were able to assess natural variation in lexical processing in children exposed to one, two, or more languages. Although our sample for the latter case is small, we believe there is some interest in reporting on them separately due to the scarcity of research on multilinguals (Unsworth, 2013, contains a review in current state of the art on language acquisition among children exposed to more than two languages).

2 Method

2.1 Participants

Results are based on data from 38 children (monolingual n = 17; bilingual n = 13; multilingual n = 8; M age = 2;8 - that is, 2 years and 8 months, range 1;11-3;4). Twenty-five additional toddlers could not be included for various reasons (parents did not return the questionnaire for 12 children; 6 were ran on a pilot version; 3 didn't want to play with the experimenters; 1 was absent on all test days; 2 had less than 30% exposure to French). The linguistic background of the toddlers was determined through a parental questionnaire that takes into account the amount of exposure and the time period during which the child learned each language. We classified as monolingual children who heard French more than 70% of the time (i.e., other languages less than 30% of the time). Bilinguals were exposed to French and exactly one other language (English n = 5; Arabic n = 3; Mandarin Chinese n = 3; Japanese n = 2; Italian n = 1; Spanish n = 1; German n = 1; Sonink n = 1), and received these 2 languages between 30% and 70% of the time. Multilinguals were children exposed to more than two languages, generally three or four with at least 30% of French exposure (English n = 6; Hebrew n = 1; Spanish n = 2; Arabic n = 2; Chinese n = 3; Polish n = 1, Italian n = 1; Tagalog n = 1; Ilocano n = 1; Russian n = 1; Romanian n = 1; Lingala n = 1; Dutch n = 1; Wolof n = 1; Creole n = 1).

2.2 Procedure and Stimuli

During the test, each child sat next to an experimenter in a quiet environment such as the library of the daycare. The test was implemented on two ipads® covered with child-friendly protective cases (Leader Price®). Since two children were tested at a time (each accompanied by one experimenter), during the main task both child-experimenter dyads wore soundproof age-appropriate headphones, connected to their tablet via a splitter inserted through a custom-made hole, to ensure accurate sound perception and minimize interaction across dyads.

The stimuli were selected using a procedure described in more detail elsewhere (Maniel, 2016). In a nutshell, there were a total of 44 pairs of words, which included nouns (n = 23 pairs), verbs (n = 12) and adjectives (n = 9) with diverse frequencies of occurrence: high (n = 17 pairs), moderate (n = 14 pairs), low (n = 13 pairs). The frequency levels were established as a function of word occurrence in two public French corpora. Prompts were adapted to the word category; for instance, for objects they were ”touche le X” *touch the X*, but for verbs ”touche celui qui X” *touch the one that Xs*. Paired visual stimuli were selected to be subjectively similar in complexity and color. Pairs were also matched on the following features: masculine/feminine; singular/plural; animals/people/objects. The touch-screen test started with 3 training trials (one from each word category, all of them high frequency; responses were excluded from the analysis) followed by 41 test ones. In each, an on-screen character with a prerecorded voice provided a prompt ”touche le X”, that matched either the left or the right picture. A correct response resulted in the on-screen character providing positive feedback (e.g. one hears ”oui, c'est ça!” *yes, thats it!* and the character jumps up and down). If the touch was to the incorrect referent, then no feedback was provided. The response time was logged from the offset of the prompt. Following the CCT, responses longer than 7s were excluded from consideration for both response time and accuracy analyses. The order of presentation of the different test pairs was pseudo-randomized to avoid having three trials of the same type in a sequence, or more than two correct responses on the same side. The child was free to stop the test early.

3 Results

Since we expected that the variances across the groups would not be equal, we decided to use non-parametric analyses. The Kruskal Wallis test was used for initial comparisons combining all three groups and, when significant, we followed up with pairwise Welch's tests. Scripts and data allowing reproducibility of these results can be downloaded from `https://osf.io/u2xyc/`.

We found no significant difference between the groups in terms of the number of trials completed and trials attempted (i.e., where a response was recorded before 7 s) by the child [$\chi^2(2) = 2.25$, p = .33], although there was a numerical trend for lower numbers of trials completed for children exposed to more than one language (bilinguals completed 31, and attempted 29 on average; multilinguals completed 31, and attempted 28) than monolinguals (completed 35, and attempted 23).

Following criteria set in advance of data analyses, only children having produced valid responses for at least 9 trials were included for the accuracy analyses (3 children were excluded). In these accuracy analyses, results aligned with predictions made on the basis of number of languages in the input, with higher percentages of correct responses for monolinguals (monolingual, M = 86.93, SD = 13.46) than bilinguals (M = 73.86, SD = 19.31), and for the former than multilinguals (M = 67.31, SD = 16.46; see Figure 1), differences that were significant in a three-way comparison [$\chi^2(2) = 8.1$, p = .02]. Pairwise comparisons involving monolinguals achieved significance in one-tailed Welch's tests [versus bilinguals, t(18.95) = 1.99, p = .03; versus multilinguals [t(12.11) = 2.89, p = .007]; but not that comparing bilinguals versus multilinguals [t(16.76) = 0.81, p = .21].

Response times were extracted from all correct trials, and the median was estimated for children having at least 4 valid responses, resulting in the same 35 children being included (see Figure 2). We report the median as a more accurate measure since distributions of reaction time are not normal. There were no significant differences between groups [$\chi^2(2) = 2.36$, p = .31], with numerically shorter responses among monolinguals (Mdn = 2121 ms, SD = 379) than bilinguals (Mdn = 2461, SD = 551; t(18.78) = 1.47; p =.08); multilinguals (Mdn = 2401, SD = 565) were not slower than bilinguals.

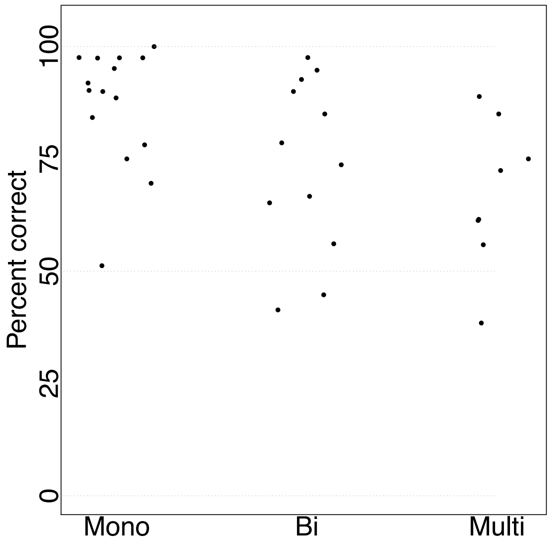

Figure 1: Accuracy as a function of language experience.

4 Discussion

We found that all three groups completed over 30 trials on average, a high number given their young age. Additionally, analyses reported elsewhere (Maniel, 2016) demonstrate that these data are sensitive to a number of individual characteristics, with strong inverse correlations between age and response time, for instance. This gave us confidence to explore the relationship between language experiences and lexical outcome measures. Our results show the same pattern as the one previous study using touch screens: Just like Poulin-Dubois et al. (2013), we find a numerical advantage whereby monolinguals have higher accuracies and faster responses than bilinguals, but the difference is not always statistically significant.

The fact that some of these differences did not reach significance could be lead to three (mutually compatible) explanations. First, laboratory-testing may lead to more sensitive measures than non-laboratory testing. Second, testing toddlers on just one of their languages could narrow group differences compared to when bilingual toddlers have to complete the same test in both languages, due to interference effects. Neither of these two explanations fit current data well since they both predict greater

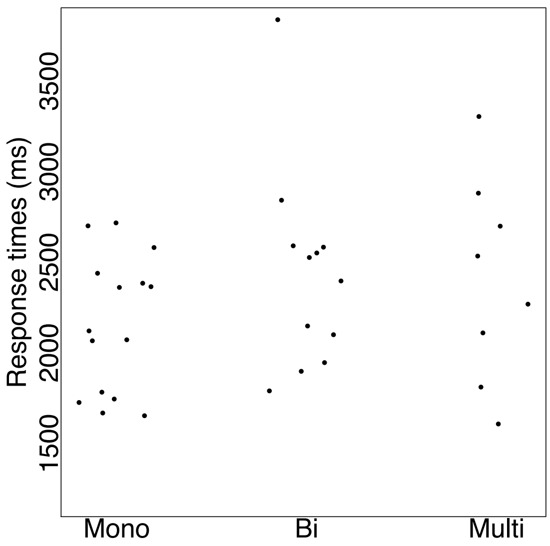

Figure 2: Response time as a function of language experience.

differences in Poulin-Dubois et al. (2013) than here because they tested bilingual children in both their languages, and they tested in the lab, unlike us – yet we obtained similar results. Third, perhaps touch-based tests are not as sensitive as other measures, because an overt motor response is required that reduces the impact of cognitive differences in processing. Only one study has been carried out comparing visual and touch-based responses, and data pertained monolingual 16-month-olds (Hendrickson et al., 2015). In this (admittedly much younger) sample, visual and haptic responses were weakly correlated (r = .4), but haptic responses were more strongly correlated with parental vocabulary reports than visual response times (r = .3 versus r = .15, respectively), lending no support to the contention that haptic responses are *less* sensitive to individual variation than visual responses (at least in the current procedure, with visual and auditory items that do not repeat, varied in form, etc.).

All this being said, we believe that it is premature, not to mention statistically inappropriate, to interpret numeric results as evidence for differences between the groups, or p-values above .05 as evidence for equality between groups. Instead, we contend that the use of such a portable method could allow more widespread testing, leading to the larger sample sizes that are necessary to capture effects that may be relatively subtle, and that are embedded in an intrinsically variable population. Indeed, we observed here that bilinguals tended to have larger standard deviations than their monolingual peers, in consonance with the expectation that the former population contains children growing up in a range of conditions. In future work, we intend to test more toddlers so as to increase our statistical power. This is crucial given that in our sample, as in previous work on the topic, we observe a wide range of variation in a number of key factors, such as the relative proportions in which the languages are present in the environment and caregivers' education level. Although in theory a bigger sample size could allow measurements on the impact of specific characteristics of the other languages on the common language (such as proportion of words that are cognates), we believe we would not be in an ideal position to investigate this given the astounding variability in terms of the other languages spoken (see Methods section). This kind of question would benefit from cross-lab collaborations studying bilingual and multilingual populations where the other languages are more stable, such as English-French in Canada or Quichua-Spanish in Argentina. Getting a better handle on such empirical variation is a pre-requisite to attempting a complete answer to key questions on lexical development, including in the case of children exposed to multiple languages.

Acknowledgments

We thank first and foremost the children and their families, without whose help this work would have been impossible. We are very grateful for the invaluable institutional support of the Direction de Familles et Petite Enfance of Paris, the three daycares who opened their doors to us, and Anne-Caroline Fievet at the LSCP Babylab. Our recognition also to our colleagues at the LSCP and CIIPME, with whom we are collaborating in related research, most saliently to Celia Rosemberg, whose Spanish stimuli we adapted in the present test. The present work was supported by the Agence Nationale pour la Recherche (ANR-14-CE30-0003 MechELex, ANR-10-IDEX-0001-02 PSL*, ANR-10-LABX-0087 IEC).

References

[Bialystok et al.2010] Ellen Bialystok, Gigi Luk, Kathleen F Peets, and Sujin Yang. 2010. Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, 13(04):525–531.

[de Houwer2009] Annick de Houwer. 2009. *Bilingual first language acquisition*. Multilingual Matters Textbook, Bristol.

[Fernald et al.1998] Anne Fernald, John P Pinto, Daniel Swingley, Amy Weinbergy, and Gerald W McRoberts. 1998. Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, 9(3):228–231.

[Friend and Keplinger2008] Margaret Friend and Melanie Keplinger. 2008. Reliability and validity of the computerized comprehension task (CCT): Data from American English and Mexican Spanish infants. *Journal of Child Language*, 35(01):77–98.

[Hendrickson et al.2015] Kristi Hendrickson, Samantha Mitsven, Diane Poulin-Dubois, Pascal Zesiger, and Margaret Friend. 2015. Looking and touching: What extant approaches reveal about the structure of early word knowledge. *Developmental Science*, 18(5):723–735.

[Hoff2013] Erika Hoff. 2013. Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, 49(1):4.

[Hurtado et al.2008] Nereyda Hurtado, Virginia A Marchman, and Anne Fernald. 2008. Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, 11(6):F31–F39.

[Maniel2016] Charlotte Maniel. 2016. Effet du statut socio-économique sur la compréhension de mots des jeunes enfants à Paris. Master's thesis, Université Pierre et Marie Curie.

[Marchman et al.2010] Virginia A Marchman, Anne Fernald, and Nereyda Hurtado. 2010. How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish–English bilinguals. *Journal of Child Language*, 37(04):817–840.

[Pearson et al.1997] Barbara Z Pearson, Sylvia C Fernández, Vanessa Lewedeg, and D Kimbrough Oller. 1997. The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, 18(01):41–58.

[Place and Hoff2011] Silvia Place and Erika Hoff. 2011. Properties of dual language exposure that influence 2-year-olds bilingual proficiency. *Child Development*, 82(6):1834–1849.

[Poulin-Dubois et al.2013] Diane Poulin-Dubois, Ellen Bialystok, Agnes Blaye, Alexandra Polonia, and Jessica Yott. 2013. Lexical access and vocabulary development in very young bilinguals. *International Journal of Bilingualism*, 17(1):57–70.

[Shneidman and Goldin-Meadow2012] Laura A Shneidman and Susan Goldin-Meadow. 2012. Language input and acquisition in a Mayan village: How important is directed speech? *Developmental Science*, 15(5):659–673.

[Unsworth2013] Sharon Unsworth. 2013. Current issues in multilingual first language acquisition. *Annual Review of Applied Linguistics*, 33:21–50.