CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Toddlers raised in multi-dialectal families learn words better in accented speech than those raised in monodialectal families

Natalia KARTUSHINA[1,2] (iD), Audun ROSSLUND[2], and Julien MAYOR[1] (iD)

[1]Department of Psychology, Faculty of Social Sciences, University of Oslo Forskningsveien 3A, Harald Schjelderups hus, 0373 Oslo, Norway and [2]Center for Multilingualism in Society across the Lifespan, Faculty of Humanities, University of Oslo, Niels Henrik Abels vei 36, 0313 Oslo, Norway Natalia.Kartushina@iln.uio.no, Audun.Rosslund@iln.uio.no, Julien.Mayor@psykologi.uio.no
Corresponding author: E-mail: Natalia.Kartushina@iln.uio.no

Abstract

Multi-accent environments offer rich but inconsistent language input, as words are produced differently across accents. The current study examined, in two experiments, whether multi-accent variability affects infants' ability to LEARN WORDS and whether toddlers' prior experience with accents modulates learning. In Experiment 1, two-and-a-half-year-old Norwegian toddlers were exposed, in their kindergarten, twice per day for one week, to a child-friendly audiovisual tablet-based e-book containing four novel pseudowords. Half of the toddlers heard the story in three Norwegian accents, whereas the other half heard it in one Norwegian accent. The results revealed no differences between conditions, suggesting that multi-accent variability did not hinder toddlers' word learning. In experiment 2, two-and-a-half-year-old Norwegian toddlers were exposed, in their homes, for one week, to the e-book featuring three Norwegian accents. The results revealed overall better learning in toddlers raised in bi-dialectal households, as compared to mono-dialectal peers – suggesting that accent exposure benefits learning in multi-accent environments.

Keywords: multi-accent; word learning; book reading; accent variability; e-book

Introduction

Many infants grow up in multi-accent environments, whereby language input is substantially variable and rich. For instance, in Oslo, 30% of the population speak Norwegian as their second language, and among the remaining 70% of speakers, ~30% use a different dialect from the one spoken in Oslo area (http://statistikkbanken.oslo.kommune.no/). Hence, many infants growing up in Oslo are naturally exposed to accent variation, which provides rich but inconsistent language

CrossMark

input, as words are produced differently across accents. For example, the word 'farm' *gård*, is realized differently depending on the speaker's dialect, with a trill or a tap /r/- /ɡoːr/ – in Eastern dialect and an uvular /ʁ/ – /ɡoːʁ/ – in Western dialect. In addition to differences in sound pronunciation, differences in the use of lexical tones between Norwegian dialects offer a challenging task for Norwegian infants exposed to dialects. While our ongoing research project examines whether accent variability affects word recognition and word comprehension in 12-month-old Norwegian infants (Kartushina & Mayor, accepted), the current study addressed whether accent variability in input affects word learning in two-and-a-half-year-old Norwegian toddlers, and examines the role of toddler's accent background (growing up in mono-dialectal vs. bi-dialectal family) in modulating the effect of multi-accent variability.

Effects of exposure to accent variability from birth

Previous research (in English- and Dutch-learning infants) has shown that multi-accent input at home affects early language development. For instance, in comparison to infants of the same age receiving uniform input, English-learning infants growing up in bi-accent families exhibit no word recognition at 14 months of age (van Heugten & Johnson, 2017), fail to detect word mispronunciations (Durrant, Delle Luche, Cattani & Floccia, 2015). and show word comprehension only in the dominant regional dialect at 20 months of age (Floccia, Delle Luche, Durrant, Butler & Goslin, 2012). These studies suggest that growing up in a bi-accent family affects the establishment of stable early word representations across accents.

Other studies, however, suggest that bi-accent toddlers 'catch up' with their mono-accent peers shortly after their second birthday. For example, bi-accent Dutch toddlers adjust their signal-to-word mapping strategies for both dialects from 24 months of age (van der Feest & Johnson, 2016) and bi-accent Canadian-English toddlers show similar word comprehension latencies as their mono-accent peers at 34 months of age (Buckler, Oczak-Arsic, Siddiqui & Johnson, 2017). Yet, word comprehension tasks, where toddlers have to recognize an orally presented familiar word and match it to one of the two pictures shown on the screen, might be ill-suited to detect peculiarities in language processing of bi-accent three-year-old toddlers, as they might show ceiling effect at this age. A recent study tested multi-accent 34-month-old Canadian-English toddlers in a demanding word recognition task (without visual cues) and revealed that speech processing difficulties in bi-accent toddlers persisted beyond the second birthday (Buckler & Johnson, 2020) Even more striking, studies in adults, using perceptual sensitivity to phonemic mispronunciation (Chen, Rattanasone, Cox & Demuth, 2017) and dialect-switching (Kirk, Kempe, Scott-Brown, Philipp & Declerck, 2018) tasks, have shown that bi-dialectal adults have more relaxed phonemic categories (as compared to mono-dialectal adults) and show processing costs when switching from one dialect to another. Overall, adults learn words faster in one accent, than in two accents, when phonological forms mismatch by one phonological feature, e.g., /vig/ = /vIg/ (Muench & Creel, 2013). Yet, multi-accent exposure appears to promote accent adaptation, as more experience with regional German accents (dialects) has been associated with better regional-accent (dialect) adaptation in nine-year-old German children (Levy, Konieczny & Hanulíková, 2019). In sum, these studies suggest that early bi-accent exposure has life-lasting effects on phonological and lexical processing of bi-accent listeners, yet it might benefit their adaptation to accent variation.

Effects of brief exposure to accent variability

It is noteworthy that while multi-accent variability appears to affect early linguistic processing in young children exposed to accents from birth, short exposure to unfamiliar accent speech seems to benefit language processing in infants with no previous multi-accent experience. For instance, brief exposure to multiple accents enabled 18-month-old English-learning infants growing up in mono-accent families to recognize words spoken in an unfamiliar accent, suggesting that multi-accent variability may support novel accents understanding (Potter & Saffran, 2017). Similarly, listening to a familiar story (so infants knew the intended word forms) read in an unfamiliar Australian English accent prior to the test enabled 15-month-old Canadian English-learning infants to recognize familiar words from nonsense words in Australian English accent (van Heugten & Johnson, 2014). Analogously, 2-minute exposure to an unfamiliar foreign(Spanish)-accent speech enabled 24-month-old mono-accent American-English infants to cope with this accent (in a word learning task) better than did their peers exposed to familiar local-accent speech (Schmale, Cristia & Seidl, 2012) or not exposed to accented speech at all (Schmale, Hollich & Seidl, 2011). These studies suggest that short (multi-)accent exposure promotes the mappings between the words in the native and unfamiliar accents, and boosts accent adaptation in infants with no prior accent experience.

Yet, it remains unclear whether 'natural' multi-accent exposure, as reflected, for example, by infants' contact with caregivers speaking different accents in a day-care centre during routine activities (e.g., storybook reading), would affect infants' ability to LEARN WORDS from this variable input and whether toddlers' accent experience (at home) modulates the effect. Previous research on minimal word pair (e.g., bih/tih) learning in 14-month-old infants has shown that infants benefit from variability in input, be it multi-speaker (Rost & McMurray, 2009, 2010) or general acoustic variability present in a one-speaker input (Galle, Apfelbaum & McMurray, 2015). These results, taken together with the results of a computational simulation study (Apfelbaum & McMurray, 2011), suggest that variability in lexically irrelevant acoustic information (speaker's pitch, amplitude, intensity, etc.) strengthens word learning by promoting associations between lexically relevant, RELATIVELY CONSTANT CUES (e.g., voice onset time to distinguish /b/ and /t/) and objects.

However, note that differences in word pronunciation between accents arise, in particular, from INCONSTANCY IN THE USE OF LEXICALLY RELEVANT cues ('bath' is produced with an /æ/ in American English and with an /ɑː/ in British English), while the phonological structure of a word typically remains preserved. Research in young infants revealed that infants fail to learn new words and display poor sound discrimination if lexically relevant cues are highly variable in ambient speech (Cristia, 2011; Rost & McMurray, 2010). These results suggest that multi-accent variability, featuring variability in lexically relevant cues, might hinder word learning. Yet, from the age of 19 months infants readily adapt to variability in lexical cues and begin to be able to accommodate unfamiliar accents (Best, Tyler, Gooding, Orlando & Quann, 2009; Mulak, Best, Tyler, Kitamura & Irwin, 2013; van Heugten & Johnson, 2014) by presumably extending (or relaxing) their phonemic categories in order to accommodate deviating examples (Schmale et al., 2012; Schmale, Seidl & Cristia, 2015) or changing the mapping between the acoustic-phonetic form and the representation (e.g., White & Aslin, 2011). Moreover, multi-accent exposure

generalizes to the understanding of other unfamiliar accents, suggesting that variability across lexically relevant cues enables toddlers to adapt to novel pronunciations by helping them to identify those cues (e.g., phonological constancy, Best et al., 2009) that are relevant for word recognition across accents (Potter & Saffran, 2017). Therefore, these studies suggest that by facilitating the retrieval of lexically invariant forms, multi-accent variability might boost lexical learning. Yet, research in non-native sound learning has shown that variability might hinder phonological acquisition in young learners, as it places significant demands on attentional and/or cognitive resources, which may present difficulties for children (Evans & Martin-Alvarez, 2016; Giannakopoulou, Brown, Clayards & Wonnacott, 2017). In sum, the current state of research provides no definite answer to the question of whether multi-accent input (featuring variability in both lexically relevant and irrelevant cues) affects word learning in young language learners.

Current study

The current study addressed the gap in knowledge on the role of multi-accent variation on early language development by assessing whether the amount of accent variability in auditory input affects word learning in two-and-a-half-year-old Norwegian toddlers, and examining if a toddler's background (mono-dialectal vs. bi-dialectal family) modulates the effect. In contrast to previous lab-based research, the current study adopted a more natural and ecologically valid approach, where toddlers were exposed, in their kindergarten groups (Experiment 1) or at home (Experiment 2), to an audiovisual storybook, presented either in three distinct Norwegian accents (Experiment 1 and 2) or in one single Norwegian accent (Experiment 1), and were assessed, individually, on how well they learnt four novel words embedded in the story plot.

To achieve our aims, we created an original child-friendly audiovisual e-book that was presented to toddlers on a tablet, twice per day, over a duration of one week. Previous research has shown that young children readily learn words from e-books. For instance, a recent study has shown that preschool-aged children (3–5 years old) learn equally well or even more from tablet-based e-books, as compared to print books (Reich, Yau & Warschauer, 2016). O'Toole and Kannass revealed that four-year-old children learn more words from an audio narration e-book, as compared to an audio-narration print book or a live e-book, suggesting that audiovisual e-books is an engaging and effective tool to learn words (O'Toole & Kannass, 2018). In the current study, the story revolves around two animal characters, a crocodile and a pig, who go fishing and, on their way to the lake, discover four visually dissimilar novel objects (see Figure 1) that have phonotactically legal, phonologically dissimilar Norwegian novel words (pseudo-words) – see Methods for details. The story was accompanied by fifteen colourful hand-painted illustrations that were specifically designed for the study (see pictures on https://osf.io/9f6cg/). Similar to previous research (O'Toole & Kannass, 2018), we used a four-alternative forced-choice picture identification task (4AFC) to assess novel word learning.

Experiment 1

To assess the role of accent-variability on word learning, two audio narration conditions were created. In the control condition, the story was recorded by three female speakers
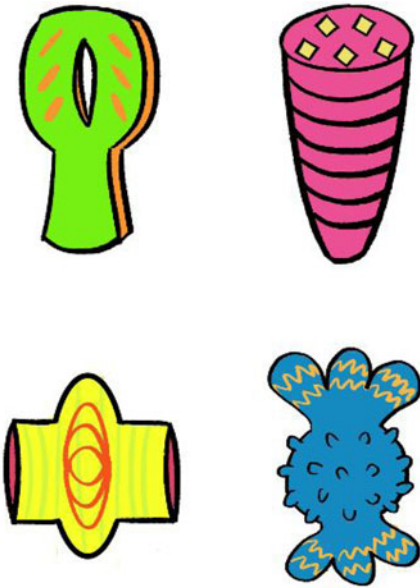
Figure 1. Novel objects (named from top left *mårku, blakko, snulle, tinkel*) used in the storybook.

in a single Norwegian accent (Eastern, Oslo), whereas in the experimental condition, the same story was recorded by the same three female speakers reading in three Norwegian accents, one per speaker (Eastern – Oslo, Western – Bergen and Northern – Tromsø). We used the term 'accent' for the sake of compatibility with previous research, yet note that each of these accents represents a different Norwegian dialect. Norwegian language can be considered as having a more complex phonology as compared to other languages examined in multi-accent infant research so far. For instance, while the majority of Indo-European languages use one acoustic cue to signal changes in vowel identity, (i.e., changes in formants, as in English 'bed' – low first formant versus 'bad' – high first formant), Norwegian, in addition, uses pitch (i.e., lexical (tone) accents, [1hendər] 'hands' vs [2hendər] 'happens') and vowel lengthening (tak [a:] – 'roof' vs takk [a] – 'thank you'). Note that although the above-mentioned three types of dialects are mutually intelligible, they are clearly recognisable even by an untrained ear for their differences at segmental (the phonetic realization of a number of sounds or their omission) and supra-segmental (e.g., the use of lexical tone accents) levels (Johnsen, 2012; Kerswill, 2016; Mæhlum & Røyneland, 2012; Røyneland, 2009). For instance, while low tone accent is used in Oslo (Eastern) dialect, high tone accent is used in Bergen (Western) dialect; that is, tone accents follow opposite patterns in Oslo and Bergen. Note that there are also differences at morpho-syntactic (e.g., differences in words' gender attribution) and lexical levels between Norwegian dialects, which are (due to the specificity of a storybook reading task) not addressed in the current study (cf. Stimuli).

If accent variability promotes word learning, then we expect that toddlers in the experimental group would learn novel words better than toddlers in the control group. However, if accent variability hinders learning, then we expect that toddlers in the experimental group would recognize fewer words than toddlers in the control group.

Method

Participants

Toddlers were recruited through an invitation sent to parents in a large municipal kindergarten in Oslo, Norway. Those parents who agreed for their child to participate in the study received, one week prior to data collection, a link to a questionnaire on nettskjema.no – a university platform for data collection – where parents signed a consent form and answered general questions on child's family and linguistic background (see folder *Questionnaires* on https://osf.io/9f6cg/). In addition, parents filled in the Norwegian version of the MacArthur-Bates Communicative Development Inventories – Words and Sentences for 18–36-month-old toddlers, to evaluate the child's productive vocabulary (Kristoffersen & Simonsen, 2012; Simonsen, Kristoffersen, Bleses, Wehberg & Jørgensen, 2014). Note that the hypotheses, methods, sample-size rationale, exclusion criteria and statistical analyses were preregistered prior to data collection on the project's OSF webpage https://osf.io/9f6cg/.

In total, we received fifty-four parental consent forms and questionnaires. After having examined the questionnaires, we excluded from the study sample 2 toddlers who had English as their native language and 5 toddlers who had reportedly less than 70% of exposure to Norwegian (our language inclusion criteria). The remaining 47 toddlers were assigned pseudo-randomly to either the experimental or the control group, making sure that toddlers' vocabulary sizes, age and gender matched between groups (Table 1). Data from 2 additional children were discarded at later steps as 1 toddler refused to be tested (he only wanted to listen to the story) and 1 toddler took part only in the first testing session, due to sickness on the other days.

Therefore, the sample consisted of 45 toddlers (24 girls), with a mean age of 2;5 (range = 2–3 years, SD = 3:7). All toddlers had Norwegian as their first language and had a minimum 70% of time exposure to Norwegian. None of the toddlers had any reported visual or auditory impairments, and none was born prematurely. All toddlers' mothers had either a Bachelor or a Master University degree (this was coincidental). Among the included 45 participants, 30 (15 in the control and 15 in the experimental group) attended all 9 listening sessions[1] and participated in all 3 identification test sessions (referred thereafter as S1, S2 and S3); the remaining 14 participants, in their majority, took part in the first two identification tests, but some of them missed the last test and/or some book exposures[2] (see Appendix A https://osf.io/9f6cg/ for details): therefore, they were not considered for the main analyses. In the final sample (n = 30), there were no group differences between participants with respect to age, gender and vocabulary size (see Table 1 for details). Note that head teachers in the kindergarten were native speakers of Norwegian, speaking Eastern Norwegian dialect. The study has been approved by the Norwegian Center for Research Data (ref. 807456). As a token of appreciation, we offered children's books for the groups that took part in our experiment.

---

[1]Two participants in the control and two participants in the experimental group missed one book reading session.

[2]Toddlers' absence on the last test was due, mainly, to parents' choice to stay home with their child or to pick their child up shortly after the lunch on Friday (note that the study was run during summer).

Table 1. Description of toddlers in the control and the experimental group.

Group	Participants	Gender	Age mean (sd)	Norwegian exposure	CDI mean (sd)
Control	n = 15	n girls = 8	28 (4)	94%	419 (183)
Experimental	n = 15	n girls = 8	30 (4)	98%	492 (168)

Note: two-tailed t-tests for age (in months), exposure and CDI revealed non-significant group differences (p>.1).

Stimuli

Visual stimuli and target novel objects

The audiovisual e-book was created specifically for the purpose of the current study, with both an original story and aquarelle hand-drawn illustrations (in a difference from previous research that used photos, e.g., Horst, Parsons & Bryan, 2011). Our aim was to make both the illustrations and the story similar to those used in the literature that children at this age group would typically be exposed to. Therefore, we used animal characters and colorful illustrations, a simple and familiar plot, and we used objects and events recurrently. The story revolved around two animal friends, a pig and a crocodile, who went fishing at a lake. On their way to the lake, the friends discovered several objects that they later would be using to catch fish with, although none of them would be good enough to attract fish. The story was divided into three logic and balanced sections with respect to their length and content, corresponding to three main events in the plot: finding objects on the way to the lake, trying to fish with different objects and a discussion of the day on the way home. The story was told in a dialogue format. The moral of the story was that in order to catch fish, one must use something edible that fish like.

Among the objects that the two characters discovered on their way to the lake, there were four unfamiliar, visually distinct objects (Figure 1). These novel objects had unique visual characteristics considered not to resemble any other familiar objects in general and those used in the storybook, in particular. These novel objects appeared three times in the story, once in each section. Unfamiliar objects had phonotactically legal, phonologically dissimilar Norwegian pseudonames: tinkel, blakko, snulle and mårku (all masculine gender). Therefore, no child was familiar with either the novel words or the objects they referred to. Production of the four words varied across the three selected Norwegian dialects in terms of the use of the lexical tone. There were no segmental differences in target word production across dialects, apart from the word mårku: the phoneme /r/ was realized as a trill or tap /r/ in Eastern dialect and as an uvular /ʁ/ in Western dialect. Note that to highlight the mapping between a novel name (presented auditorily) and a novel object (presented visually), subtle animations were applied for each novel object via DaVinci Resolve software (version 15) expanding its size in a blinking fashion when it was referred to in the story.

There were 20 separate paper drawings that we combined into 15 illustrated book pages (see Figure 2 and folder *Stimuli* on https://osf.io/9f6cg/). The illustrations were edited using Photoshop CC 2018 (version 19.0) run on Windows 10 and were combined into a movie file before being complemented with the audio recordings using DaVinci Resolve (version 15) software.
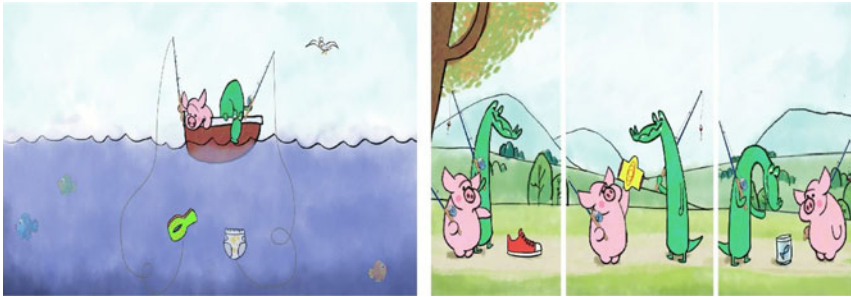
Figure 2. Examples of two illustrations/pages featuring two novel objects (mårku and snulle) from the audiovisual e-book.

Audio stimuli

The audio recordings were performed in a music recording facility, using a Røde NTG-3 condenser shotgun microphone, run through a Universal Audio Apollo Twin Duo audio interface, with Cubase Elements 9 (version 9.0.40) software on a macOS 10.14. Three female native Norwegian speakers (mean age 29 years), living in Oslo and fluent in this dialect, recorded the storybook in Oslo dialect. The speakers were instructed to imagine reading to a toddler, heightening their pitch and slowing down their tempo in a natural manner, yet maintaining dialect-specific lexical pitch, phrasal tone and morpho-syntax (cf. below). Two of these speakers (speaker 2 and 3) were not originally from Eastern Norway, but from two different parts of Norway, where Nordnorsk (North Norwegian) and Vestlandsknork (Western Norwegian) dialects are spoken, respectively. These two speakers recorded the storybook in their native dialect as well, using the same pace and child-directed speech style. Note that a word's gender varies across Norwegian dialects (a door can be masculine or feminine depending on a dialect); in order to keep the audio narration as natural as possible, within each dialect, the text was read in a dialect-congruent gender, yet all four target words were consistently masculine in all three dialects. All recordings were processed to cancel out background noise and enhance the acoustic signal. The audio recordings of speakers 2 and 3 were compared acoustically across dialects (Eastern and North Norwegian for speaker 2, and Eastern and Western Norwegian for speaker 3); the results of two-sided t-tests revealed no differences in average pitch, pitch variability or length between dialects ($p > .1$, see Appendix B https://osf. io/9f6cg/ for details on each passage and dialect). In other words, the difference between the stimuli in the experimental (three dialects) and in the control (one dialect) condition was the presence of multi-dialectal variability in the experimental condition.

In both the experimental and the control conditions, the final audiovisual e-book consisted of 16 pages (including the front page), 363 words (including the title), and it lasted approximately 4 minutes 30 seconds. To counterbalance the order of voices within and between conditions, we created 12 voice-order versions of the audio narration: 6 for the experimental and 6 for the control condition. In each voice-order version, the story was narrated by three speakers, one per story section (three in total), and their order was counter-balanced across the nine storybook exposures. For example, on the first exposure to the storybook, toddlers in both the experimental and the control group heard the first section of the story in the speaker's 1 voice, the

second section in the speaker's 2 voice and the third section in the speaker's 3 voice. On the second exposure to the storybook, toddlers in both groups heard the storybook in a different voice order: speaker 2 – speaker 3 – speaker 1, and so on. Therefore, on each storybook listening session, each novel word appeared three times and each time in a speaker-different voice.

Procedure

The study took place in a large kindergarten in Oslo and spanned over six experimental weeks. On each experimental week, 8 to 12 toddlers from different kindergarten sections were invited to take part in the study that lasted one week. Recall that parental consent forms and answered questionnaires were received a few days before each experimental week started and were used, a few days before the experiment, to create matching control and experimental groups. On each experimental week, to complete a study, a child had to participate in nine storybook listening sessions, delivered twice per day from Monday to Friday (only one storybook exposure on Friday), and three test sessions, delivered on Monday morning after the first exposure, Wednesday morning after the fifth exposure and Friday morning after the last 9th exposure to the storybook. The storybook listening sessions took place twice a day, at 10 am and 2 pm, except for the last day (Friday) that had only one exposure at 10 am.

Storybook listening sessions

Exposure to the audiovisual e-book (implemented on a Samsung Galaxy Tab S4 10.5") took place in a separate room situated relatively far from the children's groups in the kindergarten (to minimize noise interference). Except for a chair for the experimenter and pillows on the floor for the toddlers, the room was empty to avoid potential distracting elements. Eight to twelve toddlers were present for the listening sessions, and toddlers were accompanied by an educator from their own group. The study began with the experimenter greeting the children, inviting them to watch an audiovisual e-book, and ensuring all toddlers had an optimal line of sight to the tablet (between 60 and 80 cm to the tablet); the e-book then was played through the Samsung Video app, on a maximum volume capacity from the internal speaker. The order of the narrators was semi-counterbalanced across the nine exposure sessions, with the six voice-order versions alternating each other. The educators were asked by the experimenter not to interfere with the listening, unless an inattentive participant became a distractor for the rest of the group. If this was the case, the educator was asked to encourage the child to be quiet and refocus her/his attention on the story. No child was excluded from the study based on unruliness during exposure or test. In total, each storybook exposure lasted around 5 minutes.

Four-alternative forced-choice identification of novel words

The identification test was performed in the same room as the storybook listening. After the 1st, the 5th and the 9th storybook exposure session, toddlers were invited to leave the room and to go to a nearby play-area with their educators, and then, one by one, they were invited to join the experimenter in the room for the identification task. The order of toddlers was semi-counterbalanced across the sessions, with the first three toddlers tested on the first test session being the last three to be tested on the third test session. Each toddler was tested individually in a tablet-based customized four

alternative forced-choice identification task implemented on a custom-based experimental platform. On each trial, a child saw four (two by two) pictures on the tablet's screen and was prompted to touch one of them by a recorded voice diffused through the tablet's internal speaker, e.g., "Kan du ta på <target>?" ("Can you touch the <target>?). A touch response triggered the next trial. The audio instructions for the identification task were recorded, in Oslo dialect, by a female thirty-year-old native Norwegian speaker (different from those used in the book narration), who was instructed to read in a toddler-directed speech. The within-trial order of target words was randomized across toddlers. There was no timeout; yet, if a child hesitated for more than 10 seconds and did not touch the screen after an additional instruction from the experimenter, then the experimenter touched the 'Next' icon to pass to the next trial. The task consisted of four warm-up trials, followed by the test trials. Missed trials were coded as not answered 'NA' trials. The target object position was counterbalanced.

The task started with four warm-up trials, where toddlers were presented with four highly familiar objects, known by all toddlers at this age: a ball, an apple, a car and a dog (none of them was used in the storybook). In order to pass to the test trials, the toddler had to provide a correct answer in all four warm-up trials; if a toddler missed one of the trials, the experimenter started over from the beginning to ensure completion of all four warm-up trials. This precaution was necessary to make sure that all toddlers understood the task and knew how to interact with the touch screen before proceeding to the test trials.

The test trials comprised three blocks of trials (four in each): two blocks of novel-word trials and one block of familiar-word trials. In a novel-word block, on each trial, toddlers saw four novel objects from the storybook (see Figure 1) and were asked to touch one of them. In the familiar-word block, on each trial, toddlers saw four familiar objects encountered in the storybook (a diaper, a shoe, a sausage and a seagull)[3] and were asked to touch one of them. The order of blocks was fixed: first, the novel-word block, then the familiar-word block, and finally the novel-word block. In total, there were two trials for each novel word and one trial for each familiar word. There were pauses between the blocks, during which toddlers saw a smiley face on the screen and heard a recorded encouragement to go on with the task (e.g., 'Da fortsetter vi!" – 'Here we go!'). The task lasted around 3–5 minutes (depending on the toddler's speed).

The stimuli for the e-book (including hand-painted illustrations), the data and the codes are available on the project's OSF webpage https://osf.io/9f6cg/.

Results

Data processing

Participants' answers in the identification task were retrieved from the server as individual spread-sheet files and were processed automatically via a customised Matlab script: correct answers were coded as "1", erroneous answers were coded as

---

[3]These objects were not balanced in terms of number of presentation in the book, with 'shoe' being mentioned five times, 'diaper' four, and 'sausage' and 'seagull' only once. Yet, this did not affect toddlers' performance as their familiar word recognition at the first test session was at ceiling (apart from two toddlers), suggesting that toddlers knew these familiar objects used in the book before the experiment.

"0", missed trials or trials with no answers coded as "NA" were removed from the consequent analyses. Note that the statistical analyses were registered prior to the data collection on the project's OSF webpage https://osf.io/9f6cg/. We report all analyses and codes in an RMarkdown document provided in Appendix C https://osf.io/9f6cg/.

Planned analyses

A two-step procedure was applied to analyse, in *R* (R Core Team, 2012), toddlers' identification of novel words. First, we performed a generalized mixed-effect regression model (with the family = 'binomial' option) available in *lme4* package (Bates, Mächler, Bolker & Walker, 2015)[4]. The model aimed to examine (1) whether word identification improved over the three testing sessions, (2) whether there were differences in word identification between the experimental and the control group, and (3) whether the effects of session differed between the two groups. In the model, the variables group, session, gender and the group by session interaction were included as fixed factors[5]. The random structure included by-child slope (adjusted for the effect of session) and by-item intercept. Note that the initial model included a by-item slope, but it was then dropped from the analyses due to non-convergence. To provide the inference criteria for the main effects and the interactions, the *Anova* function from the *car* package was run on the model (Fox, Weisberg, Price, Adler, Bates, Baud-Bovy, Bolker, Ellison, Firth, Friendly, Gorjanc, Graves, Heiberger, Laboissiere, Maechler, Monette, Murdoch, Nilsson, Ogle, Ripley, Venables, Walker, Winsemius, Zeileis & R-Core, 2019). To compute the effect sizes for each factor, we ran a simple generalized regression model with the fixed structure including the factor of interest and the random structure identical to the one that was used in the main model, and applied the function *r.squaredGLLM* from *MuMIn* package to compute marginal and conditional R-squared for generalized mixed models, which represent the variance explained by the fixed effect and by both the fixed and the random structure, respectively (Barton, 2013). In the second step of the analysis, we examined whether children's performance on each test session was superior to chance (which was 25%). For that, each individual child's data were averaged over the eight novel word trials (four in each block) to obtain one recognition score per child and session. Given that the recognition scores were not normally distributed (p = .007, Shapiro test), two-sample Wilcoxon tests were performed for each session and group.

The results of the mixed-effect regression analysis revealed a significant effect of session only (χ^2 = 17.5, p<.001, R^2_m = 0.04, R^2_{cond} = 0.15), suggesting that more book exposure resulted in better novel word recognition (means (sds) were: S1 = 0.25 (0.19), S2 = 0.37 (0.21) and S3 = 0.46 (0.26)). There were no overall differences between the experimental (mean = 0.38, sd = 0.24) and the control (mean

---

[4]A model with an interaction between group, block and session, and an effect of gender (announced on the project's webpage) did not converge. Note that, as expected, for familiar words, the results revealed ceiling effects at each tested session and no group differences. The results for familiar words are presented in Figure 3, for the sake of comparison with novel words.

[5]Note that although our groups were matched for the vocabulary size, in line with previous research, we examined whether the vocabulary size modulated the effects of session, or the session-by-group interaction. The results revealed no effect of vocabulary or its interaction with the other factors, so we dropped this variable from the analyses.
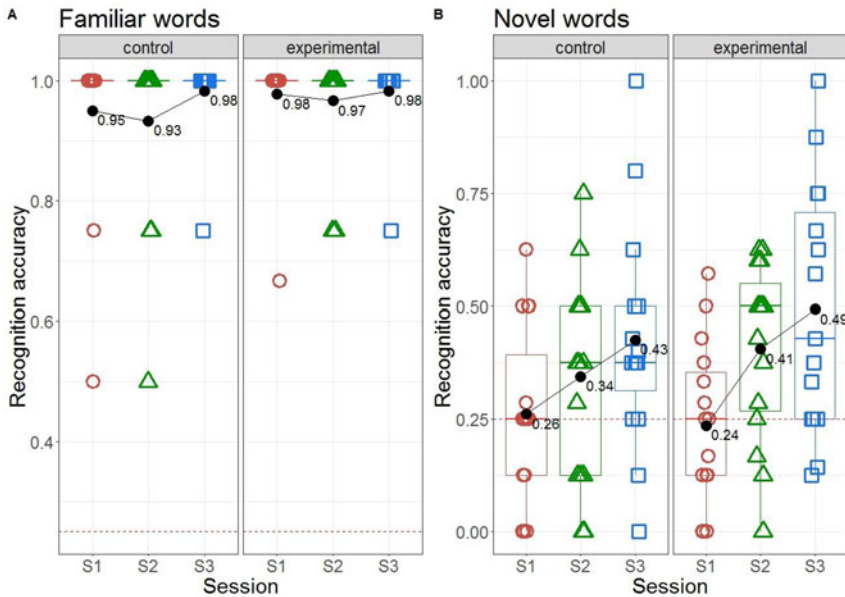
Figure 3. Recognition accuracy for familiar (A) and novel (B) words in the experimental and in the control group. Shapes indicate individual participants. Bold black dots indicate average performance. Dashed line indicates chance level.

= 0.34, sd = 0.23) group, $\chi^2 = 0.1$, p = .75; or girls (mean = 0.39, sd = 0.23) and boys (mean = 0.33, sd = 0.24), $\chi^2 = 1.1$, p = .29. The group by session interaction was also non-significant ($\chi^2 = 1.05$, p = 0.30), suggesting that both groups showed similar progress in word learning across sessions. Note, however, when toddlers' performance on each test was compared to chance, the results revealed at chance performance both in the experimental (mean = 0.24, sd = 0.18, Cohen's D = 0.08, p = .72 two-tailed) and in the control (mean = 0.26, sd = 0.20, Cohen's D = 0.08, p = .79 two-tailed) groups on S1. Yet, on S2, toddlers' performance was significantly different from chance in the experimental group only (mean = 0.41, sd = 0.20, Cohen's D = 0.79, p = .015, two-tailed test), participants in the control group performed at chance level (mean = 0.34, sd = 0.23, Cohen's D = 0.41, p = .15, two-tailed test). This result suggests that after five sessions of book exposure, toddlers in the experimental group were able to accurately recognize novel objects. Yet, the difference between the two groups on S2 was not significant (mean = 0.7, Cohen's D = 0.29, p = .19, one-tailed test). On S3, on the other hand, both groups showed accurate novel word recognition, with mean = 0.49, sd = 0.28, Cohen's d = 0.88, p = .009 (two tailed) in the experimental group and mean = 0.43, sd = 0.25, Cohen's d = 0.70, p = .02 (two tailed) in the control group.

To measure improvements in each group, we computed a difference score between each participant's average performance from S1 to S2 [S2-S1], from S2 to S3 [S3-S2] and from S1 to S3 [S3-S1]. Bigger scores indicated larger improvements. On average, after being exposed to the storybook for five times, participants in the control group improved their identification accuracy by 8%, whereas participants in the experimental group improved their identification by 17%. After additional four

exposures, both groups showed similar amounts of further improvement: 8% in the control and 9% in the experimental group. Finally, overall, as compared to performance on S1, after nine exposures to the storybook, participants in the control group improved their identification of novel words by 17%, whereas participants in the experimental group improved their identification by 25% (yet, the 8% difference was not significant, p = .18, one-tailed Wilcoxon test). To sum-up, our results suggest that dialectal variability embedded in the storybook does not hinder word learning: participants in the experimental group showed similar learning performance as participants in the control group. The numerical trend for larger benefits in the experimental group was not significant[6].

Exploratory analyses

When examining the linguistic profiles of our participants, we noticed that 9 out of 15 participants in the experimental group were exposed to two distinct Norwegian dialects at home (spoken by native speakers of Norwegian)[7], with one of the two dialects being Eastern Norwegian (Oslo dialect)[8], and the other dialect being one of the remaining two used in the e-book. To examine whether everyday dialectal exposure enabled toddlers to learn from dialectal input better as compared to those who were not exposed to dialects at home[9], we divided our participants in the experimental group in two groups as a function of a presence (yes) or an absence (no) of two (distinct) dialects at home. As can be seen in Figure 4A, participants' learning curves differed between the two groups. While both groups of toddlers appeared to improve their identification after 5 storybook exposures (S2-S1 in Figure 4B), only toddlers exposed to two dialects at home (in fact, all of them) appeared to benefit from the additional 4 storybook exposures (S3-S2 in Figure 4B); overall, toddlers who were not exposed to two dialects at home, on the other hand, showed no further improvement (just one of them showed further improvement) or even a degraded recognition. To assess this observation statistically, we performed a mixed-effect regression model analysis on the dependent measure 'improvement' (computed in previous analyses). The fixed factors were inter-session changes (S2-S1 and S3-S2), dialect at home (yes/no) and a session-by-dialect interaction; the random factor was by-child intercept. Given the exploratory nature of the analyses, the significance threshold for the p = value was set

---

[6]The analyses of the full sample (n = 44) that included participants who missed some exposures and testing sessions revealed mirroring results to those reported in the full-exposure sample; that is, (a) both groups showed similar progress in word learning across sessions, (b) participants in the experimental group showed numerically better performance at all test sessions and (c) at S2, participants in the experimental group showed numerically larger improvements (17%), than participants in the control group (5%), yet, both groups showed similar amounts of improvements (10%) at S3. Figures and detailed analyses with the statistical models and significance coefficients are in Appendices D and E.

[7]In the control group, 2 out of 15 toddlers were exposed to two dialects at home. We recall that our participants were randomly assigned to either the experimental or the control condition and any eventual reassignment of a toddler to a different group was motivated by the need to match groups for gender, age and/or vocabulary size. A thorough analysis of the linguistic profiles (and, in particular, of a dialectal status at home) was performed after having built the planned statistical models.

[8]One of the toddlers was not exposed to the community-dominant Eastern dialect at home, but to Western and Northern Norwegian dialects. When we removed this toddler from the analyses, the results remained the same.

[9]Research in older children has shown that familiarity with regional accents improves adaptation to unfamiliar accents (Levy et al., 2019).
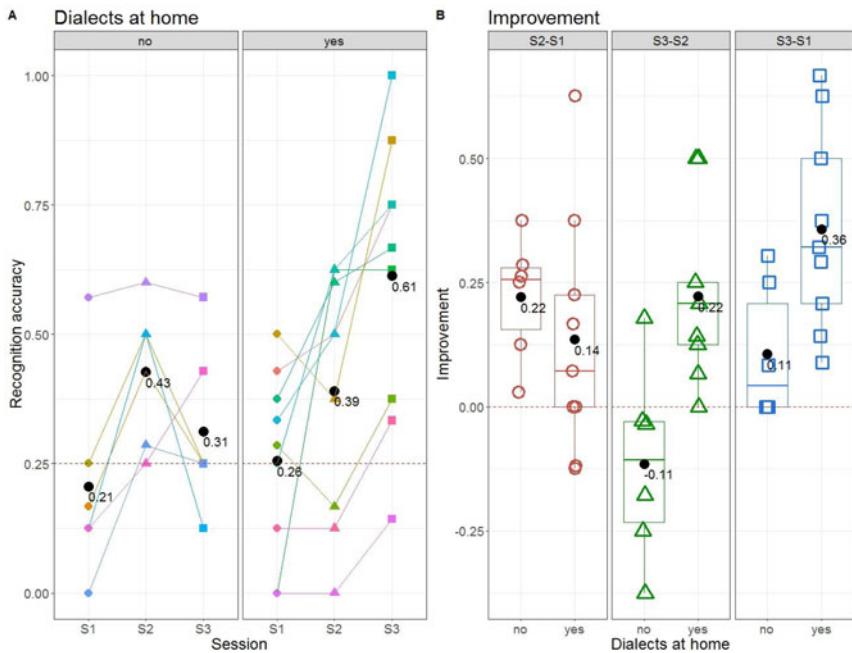
Figure 4. Recognition accuracy (A) and improvements in novel word recognition (B) in the experimental group as a function of dialectal presence (yes/no) at home. Shapes indicate individual participants. Bold black dots indicate average performance. Dashed line indicates chance level.

to 0.025 and Bayes Factor analyses were used to examine the strength of evidence presented by our data. The results revealed no main effect of inter-session (mean$_{S2-S1}$ = 0.17 (0.20), mean$_{S3-S2}$ = 0.085 (0.25), χ^2 = 1.33, p = .24) and no main effect of home dialectal exposure (with mean = 0.24 (0.22) for toddlers exposed to two dialects and mean = 0.07(0.20) for toddlers exposed to uniformly-accented input, χ^2 = 3, p = .08). However, the results revealed a significant interaction between improvement and dialect (χ^2 = 8.4, p = .0037, R^2 = 0.33). Bayes Factor analyses (Morey, Rouder & Jamil, 2015) revealed that the model with the interaction term improved the main-effects model by a factor of BF_{10} = 5.61, which presents moderate evidence in favour of the interaction effect (Lee & Wagenmakers, 2013).

A follow-up *lsmeans* test (Lenth, 2016) performed on the model revealed no difference in the amount of improvement between the two groups after 5 storybook sessions (β = 0.08, se = 0.1, t = 0.83, p = .41), but higher improvements in the bi-dialectal group after 4 additional storybook sessions (β = −0.34, se = 0.1, t = −3.28, p = .003). This indicates that, after the first 5 storybook listening sessions, toddlers exposed to two dialects at home improved their novel word recognition (mean = 0.14, sd = 0.24) to the same extent as toddlers exposed to uniform input (mean = 0.22, sd = 0.12). Yet, the 4 additional storybook listening sessions led to significantly better novel word identification (Cohen's d = 1.84) in toddlers who were exposed to two dialects at home (mean = 0.22, sd = 0.17), as compared to those who were exposed to a uniform input (mean = −0.11, sd = 0.19); the latter did not benefit from the additional four storybook listening sessions.

A comparison of the amounts of overall improvements (S3-S1) between the two groups of toddlers revealed significantly larger improvements in the bi-dialectal group (mean = 0.36, sd = 0.20), as compared to a monodialectal group (mean = 0.11, sd = 0.14, 95%CI [−0.04, 0.25]), p = .021, two-tailed, Cohen's d = 1.39, BF_{10} = 6.11. These results suggest that, by the end of the storybook-exposure experiment, among toddlers who were exposed to the story read in three distinct Norwegian dialects, those who had exposure to dialects at home learnt novel words better than those who had a uniformly-accented input at home. Note that there were no differences in the vocabulary size between the two groups of toddlers (445 and 523 words, p = .37, two-tailed) or in the ratio of boys/total (3/6 and 4/9) between the groups, suggesting that group differences in novel word identification were due to differences in previous (home) experience/familiarity with dialectal speech.

Given the incidental nature of the above-reported result (i.e., that bi-dialectal toddlers learned words better in variable multi-accent input, as compared to mono-dialectal toddlers), the (low) sample size and unbalanced distribution of toddlers between groups, we designed a follow-up experiment to test the hypothesis that growing up in bi-dialectal households can benefit word learning in multi-accent input.

Experiment 2

The current experiment stemmed from the results of the first experiment and aimed to examine the role of exposure to dialects at home on Norwegian toddlers' ability to learn words from multi-accent (dialectal) input. We used exactly the same materials as for the experimental group in Experiment 1 – that is, toddlers were exposed to the audiovisual story 'Krokko og Grynte på fisketur' that featured four novel words, presented, on each exposure session, in three distinct Norwegian dialects (one-accent condition was not used here), the order of which was counterbalanced across sessions. Then, toddlers were assessed on their learning of the novel word-object pairs, in a 4-FCI task. However, due to the Covid-19 pandemic, we were not able to test toddlers in their kindergartens, but performed data collection remotely (fully online), in their homes, via an experimental platform developed by Lo and colleagues (http://github.com/lochhh/e-Babylab) – note that the same platform was used in the original study, in the kindergarten, but it was accessed from a lab tablet. Previous research on the use of this platform, with 18–20-month-old Norwegian infants, reported no differences in infants' engagement with the task, motivation and accuracy when performing a forced-choice recognition task in the lab versus at home (Lo, Rosslund, Chai, Mayor & Kartushina, 2021). In order to maximize the number of volunteering participants and valid data points, we adapted the original testing procedure to families' everyday routines. Namely, toddlers were invited to attend to one book exposure per day (typically in the afternoon, after the kindergarten), followed by an immediate novel word 4-AFC test. The full experiment consisted of 5 sessions, each of them involved one book exposure and one identification test, run on 5 consecutive days (from Monday to Friday); therefore, on each day, parents received one link redirecting them to the day-specific task. Given the results of Experiment 1, we expected that all toddlers would learn words. Yet, if dialectal variability at home benefits word learning in multi-accent environment, then toddlers exposed to two dialects at home would learn words better than toddlers exposed to one dialect.

Methods

Participants

Monolingual Norwegian toddlers aged 28–30 months of age (n = 1055) and living in the grand area of Oslo (data available via National Registry, *Folkeregister*) were invited, by a postal mail, to take part in our 7-week on-line study run on a rolling basis between 31.08.2020 and 20.11.2020. To ensure that toddlers' age would be around 29 months while performing the tasks, invitations were sent in 7 "waves", as a function of the child's date of birth. Among 119 parents who expressed their interest to take part in the study, 56 filled in our online language background questionnaire and 42 actually completed at least one session (exposure + test, though not necessarily on the right testing days) on the platform. Given (1) inconsistencies in parents' availability to attend, for five consecutive days[10], one testing session per day with their child and (2) importance of having a homogeneous sample in terms of time intervals between testing sessions, e.g., to ensure comparable consolidation through sleep (Horváth & Plunkett, 2016), and (3) desire to maximise sample size, we considered, for the final sample, all toddlers, who attended at least 3 sessions, ran on 3 consecutive days, which forms the majority of the sample. Thirty toddlers (out of 42) successfully completed, on three consecutive days (e.g., Monday, Tuesday and Wednesday), at least 3 testing sessions; 26 of them attended 4 sessions and 24 of them attended all 5 sessions. Among the remaining 12 toddlers (out of 42), 4 attended three sessions in 1 or 2 days[11], 1 attended four sessions in 2 days, 4 attended 5 sessions in 2 or 3 days, 2 attended 2 sessions and 1 attended 1 session.

In the final sample (*n* = 30, *mean* = 28 months, ranging from 27.3 and 29.3 months, see Table 2), 14 toddlers were exposed, at home, to parents speaking one Norwegian dialect (13 of them heard Eastern dialect and 1 of them heard Southern dialect); and 16 toddlers were exposed to two Norwegian dialects, for 15 of them, Eastern (community-dominant) dialect was one of the spoken dialects at home (one toddler was exposed to Northern and South-Western dialects). Therefore, all bi-dialectal toddlers were familiar with two out of the three accents used in the e-book.

Procedure

After accepting the invitation to take part in the study, parents were invited to fill-in a language background questionnaire (the same as in Experiment 1) and were assigned to the closest experimental "wave", starting on the following Monday. A couple of days before the start of the experiment, parents received detailed instructions about the course of the experiment, indicating, among others, that they would receive one link

---

[10]Despite the instructions to run one task per day, for 5 consecutive days, many parents were not able to follow these guidelines and contributed to 2 or 3 (daily) tasks only. Among those who did perform 5 tasks, only 17 performed them on 5 consecutive days, the others, mainly those who missed one day-matching task, performed the task on the next day (or the day after), together with a day-specific/matching task. Therefore, the two tasks would be immediately following each other and not separated by night sleep, which would ensure learning consolidation (Horváth & Plunkett, 2016). In addition, performance in a word identification test after being exposed to the story for two times in a row, might be qualitatively incomparable to when word learning was assessed after one exposure only.

[11]Day-specific links to the platform were active for the whole week; some parents, who were not able to run a task on a designated day, decided to 'catch' up with the task and ran several links (2 and more sessions) in one day. Those children, who performed two tests in one day were excluded from the analyses.

Table 2. Description of toddlers growing up in monodialectal and bi-dialectal households.

Group	Participants	Gender	Age mean (sd)	Norwegian exposure	CDI mean (sd)
Monodialectal	n = 14	n girls = 9	28 (0.3)	94%	530 (94)
Bi-dialectal	n = 16	n girls = 14	28 (0.4)	98%	554 (126)

Note: two-tailed t-tests for exposure and CDI (when using both row score and percentiles) revealed non-significant group differences (p>.1).

(to the platform) per day, to be open on a touch-screen device that they have at home in order to perform a day-specific session with their child. Parents received, in the morning, one day-specific link. There were 5 links in total, one for each session. After clicking on the link, parents were directed to the platform, where they had to indicate, first, their participant identifier and then to consent not to interfere with the task and child's performance during the task. Then, an audio recorded prompt with a smiley face on the screen invited a child to watch and listen to the story 'Krokko og grynte på fisketur'. Immediately after the audiobook exposure, a smiley face would reappear on the screen inviting the child to play. As soon as the parent would click on the 'next' button, the 4-AFC task (identical to the one used in Experiment 1) would start. At the end of the session, that typically lasted between 6 and 10 minutes, a smiley face would reappear on the screen and would thank the child for the participation. There were breaks inserted between the audiobook and the test, that parents could opt to take. In order to counterbalance the order of dialects in the story across the 5 audiobook exposures (as it has been done in Experiment 1), we created 5 order lists that rotated across the testing weeks. At the end of the experiment, all parents received a gift card for a book shop.

Results

Data processing

Similar to the data in Experiment 1, participants' answers in the identification task were processed automatically via a customised Matlab script: correct answers were coded as "1", erroneous answers were coded as "0", missed trials or trials with no answers coded as "NA" were removed from the consequent analyses. Detailed outcome of an RMarkdown document is provided in Appendix F https://osf.io/9f6cg/.

Planned analyses

To examine the role of home exposure to dialects on novel word learning in multi-dialect environment, we performed generalised mixed-effect regression analysis (package afex, Singmann, Bolker, Westfall, Aust, Ben-Shachar, Højsgaard, Fox, Lawrence, Mertens, Love, Lenth & Christensen, 2021) on the dependant binomial variable accuracy with the fixed factors session (from 1 to 3, as these were attended, on 3 consecutive days, by all 30 toddlers in the sample), dialects at home (yes vs. no) and an interaction between session and dialects at home; we included toddlers' vocabulary size as a covariate. The random structure included by-subject intercept and by-item slopes (including an interaction with the main effect of session). Similar

to the analyses in Experiment 1, the *anova* function (Fox et al., 2019) was used to assess the significance of the main effects and the interactions; and r.*squaredGLMM* from *MunMin* package (Barton, 2013) was used to compute effect sizes. The results revealed a significant effect of session ($\chi^2 = 8.51$, p = .0035, $R^2_m = 0.02$, $R^2_{cond} = 0.13$), dialects at home ($\chi^2 = 5.8$, p = .015, $R^2_m = 0.05$, $R^2_{cond} = 0.14$), with bi-dialectal toddlers displaying better word recognition (m = 0.40, SD = 0.25) than mono-dialectal toddlers (m = 0.27, SD = 0.22), Cohen's d = 0.77; and a marginally significant interaction between session and dialects at home ($\chi^2 = 3.18$, p = .074) – see Figure 5. There was no effect of vocabulary size ($\chi^2 = 1.24$, p = .26)[12]. Note that the main effect of session ($\chi^2 = 9.7941$, p = 0.0017) and dialect at home ($\chi^2 = 4.49$, p = 0.034) remained similarly significant when the model was run on all 5 sessions (with 30 participants in S1-S3, 26 in S4 and 24 in S5) – see Appendix F (https://osf.io/9f6cg/) for details.

Follow-up analyses, via *lsmeans* on the model, revealed higher recognition accuracy in the bi-dialectal group at S1 ($\beta = -0.86$, $SE = 0.430$, z = -1.99, p = .045, Cohen's d = 0.31) and a marginally higher recognition at S2 ($\beta = -0.72$, SE = 0.38, z = -1.88, p = .06, Cohen's d = 0.29), yet no between-group differences at S3 ($\beta = -0.045$, SE = 0.36, z = -0.13, p = .89). Note, however, when toddlers' performance was compared to chance for each session, via one-sample two-tailed Wilcoxon signed rank test, the results revealed at chance performance at S1 for both bi-dialectal (m = 0.32, SD = 0.24, p = .26) and mono-dialectal (m = 0.17, SD = 0.16, p=.99) groups, and a significant word recognition at S2 for the bi-dialectal group only (m = 0.45, SD = 0.25, p=.0023, Cohen's d = 0.38), the mono-dialectal group still showed no word recognition (m = 0.26, SD = 0.22, p = .47). However, both groups showed significant and similar word recognition at S3, with an average recognition scores 0.42 and 0.40 in the bi-dialectal (SD = 0.25, p = .0009, Cohen's d = 0.34) and the mono-dialectal (SD = 0.22, p = .0035, Cohen's d = 0.33) group, respectively.

Discussion

The current study examined (1) whether accent variability in auditory input affects word learning in Norwegian toddlers exposed to an audiovisual e-book, and (2) whether toddlers' experience with dialects at home modulates the effect. In Experiment 1, two groups of monolingual Norwegian two-and-a-half-year toddlers, matched for vocabulary, gender and age, were exposed, nine times over a week, to an audiovisual child-friendly storybook that introduced four novel named objects. In the control group, participants heard the story in one Norwegian (Eastern) accent, recorded by three Norwegian female speakers, whereas in the experimental group the story was delivered in three Norwegian accents (Northern, Western and Eastern), recorded by the same three speakers. Voice quality (average pitch, pitch variation and intensity), as well as sentence duration, were matched between the two groups. Word learning was assessed, after the first, the middle and the last exposure, in a 4-AFC tablet-based task, where, on each trial, toddlers saw four novel objects (from the book) and were prompted to identify (touch) one of them (e.g., 'Where is the <target>?). In Experiment 2, conducted online in toddlers' homes, Norwegian

---

[12]Two of the toddlers were not exposed to the community-dominant Eastern dialect at home, but to Western and Sothern Norwegian dialects. When we removed these toddlers from the analyses, the results remained the same.
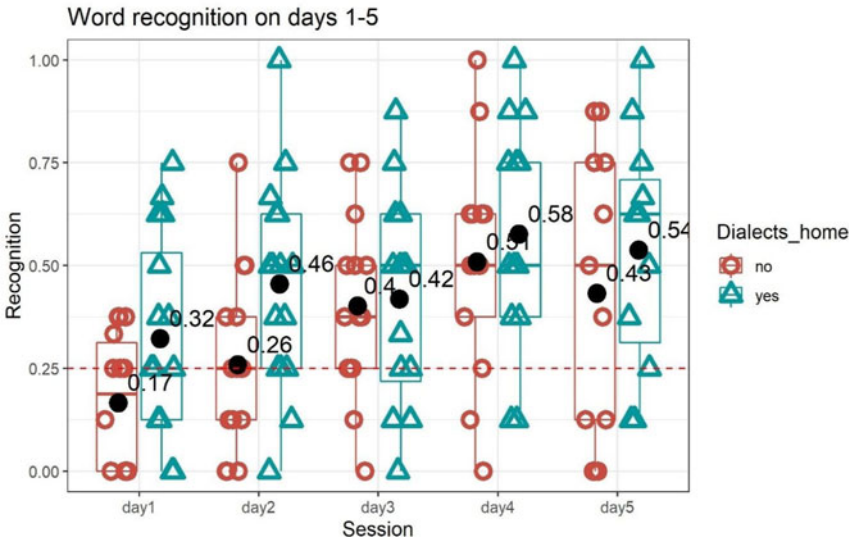
Figure 5. Recognition accuracy for novel words in toddlers exposed to dialects at home and toddlers growing up in uniformly accented households. Shapes indicate individual participants (n = 30). Bold black dots indicate average performance. Dashed line indicates chance level. Note that all 30 toddlers successfully completed the first three sessions on three consecutive days; yet only 26 and 24 toddlers completed 4 and 5 days of the study, respectively, although the last two sessions were not systematically completed on two consecutive days.

two-and-a-half-year mono- and bi-dialectal toddlers were exposed, on five consecutive days (from Monday to Friday), to the same audiovisual e-book, presented in three Norwegian accents, and were assessed – in the same 4-AFC task and after each e-book exposure – on their word retention.

Role of accent-variability on word learning

The results of Experiment 1 revealed that accent variability embedded in the storybook did not hinder (or promote, yet see the discussion below) word learning in two-and-a-half-year old Norwegian toddlers: participants in the experimental group showed similar learning performance to participants in the control group. These results suggest that toddlers adapted unfamiliar multiple accents relatively easily (as they were not delayed in learning with respect to the control group) and were able to learn words from an audio input that featured variability in lexically relevant cues – that is, differences in word pronunciation across accents. Previous research has shown that toddlers begin to be able to accommodate unfamiliar accents from the age of 19 months, as revealed by their recognition of familiar words in unfamiliar-accent speech (Best et al., 2009; Mulak et al., 2013; van Heugten & Johnson, 2014). The current study suggests that, by the age of 2,5 years, toddlers can establish abstract, accent-independent phonological representations for novel words that were produced inconsistently across accents during the exposure period and to generalize this knowledge to a community spoken accent, as the identification test was performed in Oslo (Eastern) accent.

Surprisingly, some results (i.e., overall numerically larger improvements and significant word identification already at the second 4-AFC test in the experimental group) suggest that (multi-)accent variability may have boosted word learning (Potter & Saffran, 2017; Schmale et al., 2012), in particular after the first five exposures to the e-book. Yet, the lack of a statistically significant difference between groups and the presence of an important inter-speaker variability precludes us from making this conclusion and calls for more research, in particular, with older children, who might display more homogeneous learning patterns. In sum, the analyses of two samples of participants (n = 30, who took part in all exposure sessions and n = 45, who missed some of the exposure sessions) revealed a consistent result: by the end of the study period, both groups of toddlers, i.e., those who were exposed to the audiovisual story read in one accent and in three accents, learnt words equally well – suggesting that multi-accent variability in input did not hinder word learning.

Role of familiarity with accent-variability on word learning in multi-accent input

The results of the exploratory analyses in the experimental group revealed that toddlers exposed to dialects at home learnt words significantly better than toddlers who grew up in uniformly-accented families. The results of the Bayesian analyses indicate moderate evidence ($BF_{10} = 6.11$) in favour of this large (Cohen's d = 1.39) effect. However, given that this result stemmed from an exploratory analysis and that the sample was small and unbalanced (more bi-dialectal toddlers than mono-dialectal), we designed a follow-up experiment (that was implemented online given the Covid-19 pandemic) to test the hypothesis that growing up in bi-dialectal households can benefit word learning from multi-accent input. The results of Experiment 2 provided convincing and confirmatory evidence that toddlers exposed to two dialects at home learn words better in multi-accent input than toddlers raised in uniformly accented households. This outcome confirms the results of previous studies on accent exposure, showing beneficial effects of SHORT-TERM multi-accent exposure on unfamiliar accent processing in MONO-DIALECTAL CHILDREN (Levy et al., 2019; Potter & Saffran, 2017; Schmale et al., 2012, 2015), and extends them by showing that LONG-TERM EXPOSURE to bi-dialectal speech facilitates WORD LEARNING in multi-accent environment.

The lack of differences in vocabulary size between the two groups of toddlers (and in the ratio of boys/girls in the sample) suggests that the effect is driven by toddlers' previous (home) experience/familiarity with accent variability in speech. However, given that two of the three accents used in the e-book were present (in different combinations) in the home environment of the majority of bi-dialectal toddlers, better word learning in bi-dialectal group can also be attributed to better familiarity and/or previous experience with the accents used in the e-book. Future research clearly distinguishing familiarity with specific accents versus everyday experience with multi-accent speech (growing up in bi-dialectal environment) is needed in order to better understand the role of long-term multi-accent exposure on word learning.

Although the design of our study does now allow addressing the mechanisms underlying bi-dialectal toddlers' word learning in multi-accent speech, two facilitating mechanisms can be considered. First, previous research has shown that children growing up in bi-accent families might have more relaxed phonological representations for words (Durrant et al., 2015; Floccia et al., 2012; Levy et al., 2019;

van Heugten & Johnson, 2017), which results in their acceptance of mispronounced words as real words. This 'tolerance' to deviations in word pronunciation might facilitate word recognition and comprehension across accents and contribute to a creation of abstract (common) word representations. Alternatively, bi-dialectal toddlers' advantage in word learning could be due to their prior practice learning words from dialectally variable input. In line with this interpretation, recent research with nine-year-old children has shown that more experience with (or exposure to) regional accents (dialects) is associated with better understanding of novel accents (Levy et al., 2019). To sum-up, the current state of research suggests that growing up in a bi-dialectal household affects processing of familiar accents, as bi-dialectal toddlers show less robust word recognition as compared to mono-dialectal peers; yet, it benefits word learning in multi-accent environment.

To the best of our knowledge, the current study is the first to reveal advantages of bi-dialectal home exposure on bi-dialectal toddlers' language development in a close-to-natural experimental setup. This result has practical implications, relevant for the Norwegian educational context, as it may suggest that having educators speaking different accents/dialects in a kindergarten – which is not infrequent in Norway, in particular in big cities – might enhance toddlers' ability to learn words in a multi-accent environment[13], to which they are likely be exposed progressively more in their older childhood and adolescence.

Detailed analyses of toddlers' performance at each test in Experiment 1 and 2 revealed that both mono and bi-dialectal toddlers showed stable word learning on days 1–3; yet, little progress has been observed between the last two days of exposure, with mono-dialectal toddlers showing a plateau in performance. These results suggest that toddlers growing up in mono-dialectal families, similar to bi-dialectal toddlers, may have accommodated regional accents (an ability they begin to display from the age of 19 months) and that, after several exposures to the e-book, they were able to extract novel words from multi-accent speech and to retain them. Yet, further exposure to multi-accent input resulted in little improvement in their performance. Other studies showed that variability might hinder phonological acquisition in young learners, as it places significant demands on attentional and/or cognitive resources, which are limited in children (Evans & Martin-Alvarez, 2016; Giannakopoulou et al., 2017). However, relatively little improvement in the last learning days in both mono-dialectal and bi-dialectal toddlers might also be attributed to a general cumulated fatigue observed by the end of the week, as in both experiments the last day of exposure was administered on Friday. More research, targeting toddlers with different linguistic backgrounds and potentially at different periods of language development, is needed to increase our understanding of the role of multi-accent (long-lasting) exposure on early language development. In sum, although toddlers raised in mono-dialectal families begin to be able to accommodate unfamiliar accents from the age of 19 months (they recognize words in unfamiliar

---

[13]Given the imbalance in the number of monodialectal and bidialectal toddlers in the control group, we were not able to compare their learning outcomes. This means that we cannot rule out a possibility that the advantage for multi-dialect children in the experimental group does not stem from their ability to cope better with accent variability, but rather is attributed to their general capacity to learn words (in which case, the home dialect exposure would facilitate performance in both groups). Although research in older children suggest that dialectal experience enhance the ability to cope better with accent input (Levy et al., 2019), the alternative explanation needs further investigation (currently examined in our lab).

accents and generalize knowledge to other speakers and accents), ten months later they experience more difficulties in learning words in multi-accent input – as compared to toddlers raised in bi-dialectal families – suggesting that multi-accent variability might pose initial difficulties for word learning in toddlers with no prior accent experience; yet they seem to 'catch up' with their bi-dialectal peers after being familiarized with the story and/or the accents.

Finally, additional analyses of the learning gains per word (cf. Appendix F https://osf.io/9f6cg/) revealed that, in both experiments, after 5 days of exposure, the pseudo-word *mårku* [morkɯ] was learned numerically better (e.g., in Experiment 2, there were 0.32 units of improvement in word identification: from 32% at S1 to 61% at S5), as compared to the other three novel pseudo-words used in the study (varying in improvements from 0.14 to 0.24, i.e., from 22%–30% at S1 to 37%–48% at S5). Recall, that, in addition to tone variation present in the production of the four novel words across the three accents, the word *mårku* triggers segmental variation, as the phoneme /r/ is realized as a trill or tap /r/ in Eastern dialect and as an uvular /ʁ/ in Western dialect. Larger improvements for *mårku* are in line with previous research on word learning in 14-month-old infants (Apfelbaum & McMurray, 2011; Rost & McMurray, 2009, 2010), suggesting that increased variability in lexically irrelevant cues (as differences in the realization of /r/ across accents) might boost word learning by attracting infants' attention to lexically more relevant cues. However, more research (carefully controlling for the amount of cross-accent variability in target words) is needed to assess the robustness and empirical validity of this observation.

Learning words from audiovisual e-books

Overall, our results add to a growing body of research showing that audiovisual e-books can be used as a pedagogical tool to teach novel words to children (O'Toole & Kannass, 2018; Reich et al., 2016). Moreover, our results extend the results of previous studies to a younger age group by showing that audiovisual e-books can be used effectively – even with toddlers of younger ages than previously thought (three years); as the youngest toddler in our sample was 24 months of age. Note that toddlers' improvements in word learning were not related to their vocabulary size (as reported by parents, using CDIs, Kristoffersen & Simonsen, 2012; Simonsen et al., 2014), suggesting that the content of the book was accessible to all toddlers, including those who had low vocabulary size. Yet a recent study on word learning via AUDIOVISUAL E-BOOKS (O'Toole & Kannass, 2018) revealed that children's vocabulary score, tested with the Peabody Picture Vocabulary Test (Dunn & Dunn, 2007), did explain 14% of the variance in word learning – suggesting that lexically advanced four-year-olds retrieved and memorized more words from the audiovisual e-book than toddlers with lower receptive vocabulary. Between-studies difference in the effect of vocabulary size on word learning can be potentially attributed to a simpler story plot and a fewer number of novel words in the current study as compared to O'Toole and Kannass' study (4 vs. 10). It is reasonable to expect that as toddlers grow up, they become exposed to linguistically more complex narrations, and their story understanding and word learning might become progressively more dependent on the size of their receptive vocabulary. Previous research examining the role of vocabulary size on word learning from AUDIO BOOKS in four-year-olds similarly reported larger improvements in children with higher vocabulary sizes (O'Toole & Kannass, 2018;

Sénéchal, Thomas & Monker, 1995; Strouse & Ganea, 2016). Note that children with low vocabulary scores surpass children with high vocabulary scores on learning elaborated words, suggesting that elaboration facilitates word learning in children with low vocabularies (Justice, Meier & Walpole, 2005). In sum, our results suggest that (simple) audiovisual e-books can be effectively used to learn novel words in two-and-a-half-year-old toddlers with varied receptive vocabulary sizes.

Although toddlers showed significant gains in word learning by the end of the experiment, their poor (at chance) performance right after the first exposure to the audiovisual e-book (the target words appeared three times) suggests that toddlers might have experienced difficulties when trying to perform simultaneously three tasks: processing a visual scene, attending to an audio story/plot and learning new words. Previous research on word-object associative learning, where the target object is presented alone (or with one distractor) and is clearly matched with an audio label, revealed that 28–36-month-old toddlers (approximately of the same age as the toddlers tested in the current study) successfully learnt four novel words after five target-word repetitions (Ackermann, Lo, Mani & Mayor, 2020). These results suggest that learning words from an audiovisual e-book might require more exposure to/ repetitions of the stimuli, as compared to a simple object labelling; this observation aligns with previous research showing that repetition promotes word learning from storybooks (Horst et al., 2011).

Conclusions

The current study examined whether accent variability in auditory input affects word learning in two-and-a-half-year Norwegian toddlers exposed to an audiovisual tablet-based storybook and whether toddlers' experience with dialects at home modulates the effect. The results of Experiment 1 revealed that multi-accent variability embedded in the audio narration did not hinder toddlers' learning of novel words introduced in the book, as they learnt similarly well as toddlers who heard the book in one Norwegian accent. Yet, the results of Experiment 2 revealed that, in the multi-accent audio condition, toddlers exposed to dialects at home showed significantly larger improvements than toddlers growing up in uniformly-accented families, suggesting that dialectal input at home benefits learning in multi-dialectal environments.

References

Ackermann, L., Lo, C. H., Mani, N., & Mayor, J. (2020). Word learning from a tablet app: Toddlers perform better in a passive context. *PLOS ONE*, *15*(12), e0240519. https://doi.org/10.1371/journal. pone.0240519

Apfelbaum, K. S., & McMurray, B. (2011). Using Variability to Guide Dimensional Weighting: Associative Mechanisms in Early Word Learning. *Cognitive Science*, *35*(6), 1105–1138. https://doi.org/10.1111/j. 1551-6709.2011.01181.x

Barton, K. (2013). Package 'MuMIn.' *Version*, *1*, 18.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Best, C. T., Tyler, M. D., Gooding, T. N., Orlando, C. B., & Quann, C. A. (2009). Development of Phonological Constancy. *Psychological Science*, *20*(5), 539–542. https://doi.org/10.1111/j.1467-9280.2009.02327.x

Buckler, H., & Johnson, E. K. (2020). *Routine exposure to accent variation has a lasting effect on young children's word recognition abilities*. ICIS.

Buckler, H., Oczak-Arsic, S., Siddiqui, N., & Johnson, E. K. (2017). Input matters: Speed of word recognition in 2-year-olds exposed to multiple accents. *Journal of Experimental Child Psychology*, *164*, 87–100. https://doi.org/10.1016/j.jecp.2017.06.017

Chen, H., Rattanasone, N. X., Cox, F., & Demuth, K. (2017). Effect of early dialectal exposure on adult perception of phonemic vowel length. *Journal of the Acoustical Society of America*, *142*(3). https://doi.org/10.1121/1.4995994

Cristia, A. (2011). Fine-grained variation in caregivers' /s/ predicts their infants' /s/ category. *The Journal of the Acoustical Society of America*, *129*(5), 3271–3280. https://doi.org/10.1121/1.3562562

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4 Peabody Picture Vocabulary Test 4th Edition*. https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Academic-Learning/Brief/Peabody-Picture-Vocabulary-Test-%7C-Fourth-Edition/p/100000501.html

Durrant, S., Delle Luche, C., Cattani, A., & Floccia, C. (2015). Monodialectal and multidialectal infants' representation of familiar words. *Journal of Child Language*, *42*(02), 447–465. https://doi.org/10.1017/S0305000914000063

Evans, B. G., & Martin-Alvarez, L. (2016). *Age-related differences in second-language learning? A comparison of high and low variability perceptual training for the acquisition of English /i/-/I/ by Spanish adults and children*. International Symposium on the Acquisition of Second Language Speech, New Sounds, Aarus, Denmark.

Floccia, C., Delle Luche, C., Durrant, S., Butler, J., & Goslin, J. (2012). Parent or community: Where do 20-month-olds exposed to two accents acquire their representation of words? *Cognition*, *124*(1), 95–100. https://doi.org/10.1016/j.cognition.2012.03.011

Fox, J., Weisberg, S., Price, B., Adler, D., Bates, D., Baud-Bovy, G., Bolker, B., Ellison, S., Firth, D., Friendly, M., Gorjanc, G., Graves, S., Heiberger, R., Laboissiere, R., Maechler, M., Monette, G., Murdoch, D., Nilsson, H., Ogle, D., Ripley, B., Venables, W., Walker, S., Winsemius, D., Zeileis, A., & R-Core. (2019). *car: Companion to Applied Regression* (3.0-3) [Computer software]. https://CRAN.R-project.org/package=car

Galle, M. E., Apfelbaum, K. S., & McMurray, B. (2015). The Role of Single Talker Acoustic Variation in Early Word Learning. *Language Learning and Development : The Official Journal of the Society for Language Development*, *11*(1), 66–79. https://doi.org/10.1080/15475441.2014.895249

Giannakopoulou, A., Brown, H., Clayards, M., & Wonnacott, E. (2017). High or low? Comparing high and low-variability phonetic training in adult and child second language learners. *PeerJ*, *5*. https://doi.org/10.7717/peerj.3209

Horst, J. S., Parsons, K. L., & Bryan, N. M. (2011). Get the Story Straight: Contextual Repetition Promotes Word Learning from Storybooks. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00017

Horváth, K., & Plunkett, K. (2016). Frequent daytime naps predict vocabulary growth in early childhood. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *57*(9), 1008–1017. https://doi.org/10.1111/jcpp.12583

Johnsen, S. S. (2012). A diachronic account of phonological unnaturalness. *Phonology*, *29*(3), 505–531. https://doi.org/10.1017/S0952675712000243

Justice, L. M., Meier, J., & Walpole, S. (2005). Learning new words from storybooks: An efficacy study with at-risk kindergartners. *Language, Speech, and Hearing Services in Schools*, *36*(1), 17–32.

Kartushina, N., & Mayor, J. (accepted pending data collection). Coping with dialects from birth: Role of variability on infants' early language development. Insights from Norwegian dialects. *Developmental Science*

Kerswill, P. (2016). Brit Mæhlum and Unn Røyneland: Det norske dialektlandskapet. *Norsk Lingvistisk Tidsskrift*, *34*(1). http://ojs.novus.no/index.php/NLT/article/view/1295

Kirk, N. W., Kempe, V., Scott-Brown, K. C., Philipp, A., & Declerck, M. (2018). Can monolinguals be like bilinguals? Evidence from dialect switching. *Cognition*, *170*, 164–178. https://doi.org/10.1016/j.cognition.2017.10.001

Kristoffersen, K. E., & Simonsen, H. G. (2012). *Tidlig språkutvikling hos norske barn. MacArthur-Bates foreldrerapport for kommunikativ utvikling*. Novus.

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press.

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, *69*(1), 1–33. https://doi.org/doi:10.18637/jss.v069.i01

Levy, H., Konieczny, L., & Hanulíková, A. (2019). Processing of unfamiliar accents in monolingual and bilingual children: Effects of type and amount of accent experience. *Journal of Child Language*, 1–25. https://doi.org/10.1017/S030500091800051X

Lo, C. H., Rosslund, A., Chai, J. H., Mayor, J., & Kartushina, N. (2021). Tablet assessment of word comprehension reveals coarse word representations in 18–20-month-old toddlers. *Infancy*, *26*(4), 596–616. https://doi.org/10.1111/infa.12401

Mæhlum, B., & Røyneland, U. (2012). *Det norske dialektlandskapet: Innføring i studiet av dialekter*. Cappelen Damm akademisk.

Morey, R. D., Rouder, J. N., & Jamil, T. (2015). BayesFactor: Computation of Bayes factors for common designs. *R Package Version* 0.9, *9*, 2014.

Muench, K. L., & Creel, S. C. (2013). Gradient phonological inconsistency affects vocabulary learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(5), 1585–1600. https://doi.org/10.1037/a0032862

Mulak, K. E., Best, C. T., Tyler, M. D., Kitamura, C., & Irwin, J. R. (2013). Development of phonological constancy: 19-month-olds, but not 15-month-olds, identify words in a non-native regional accent. *Child Development*, *84*(6), 2064–2078. https://doi.org/10.1111/cdev.12087

O'Toole, K. J., & Kannass, K. N. (2018). Emergent literacy in print and electronic contexts: The influence of book type, narration source, and attention. *Journal of Experimental Child Psychology*, *173*, 100–115. https://doi.org/10.1016/j.jecp.2018.03.013

Potter, C. E., & Saffran, J. R. (2017). Exposure to multiple accents supports infants' understanding of novel accents. *Cognition*, *166*, 67–72. https://doi.org/10.1016/j.cognition.2017.05.031

R Core Team. (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org.html

Reich, S., Yau, J., & Warschauer, M. (2016). Tablet-Based eBooks for Young Children: What Does the Research Say? *Journal of Developmental & Behavioral Pediatrics*, *37*(7), 585–591. https://doi.org/10.1097/DBP.0000000000000335

Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science*, *12*(2), 339–349. https://doi.org/10.1111/j.1467-7687.2008.00786.x

Rost, G. C., & McMurray, B. (2010). Finding the Signal by Adding Noise: The Role of Noncontrastive Phonetic Variability in Early Word Learning: Phonetic Variability and Word Learning. *Infancy*, *15*(6), 608–635. https://doi.org/10.1111/j.1532-7078.2010.00033.x

Røyneland, U. (2009). Dialects in Norway: Catching up with the rest of Europe? *International Journal of the Sociology of Language*, *2009*(196–197), 7–30. https://doi.org/10.1515/IJSL.2009.015

Schmale, R., Hollich, G., & Seidl, A. (2011). Contending with foreign accent in early word learning. *Journal of Child Language*, *38*(05), 1096–1108. https://doi.org/10.1017/S0305000910000619

Schmale, R., Cristia, A., & Seidl, A. (2012). Toddlers recognize words in an unfamiliar accent after brief exposure: Brief exposure to an unfamiliar accent. *Developmental Science*, *15*(6), 732–738. https://doi.org/10.1111/j.1467-7687.2012.01175.x

Schmale, R., Seidl, A., & Cristia, A. (2015). Mechanisms underlying accent accommodation in early word learning: Evidence for general expansion. *Developmental Science*, *18*(4), 664–670. https://doi.org/10.1111/desc.12244

Sénéchal, M., Thomas, E., & Monker, J.-A. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology*, *87*(2), 218. https://doi.org/10.1037/0022-0663.87.2.218

Simonsen, H. G., Kristoffersen, K. E., Bleses, D., Wehberg, S., & Jørgensen, R. N. (2014). The Norwegian Communicative Development Inventories: Reliability, main developmental trends and gender differences. *First Language*, *34*(1), 3–23.

Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M. S., Højsgaard, S., Fox, J., Lawrence, M. A., Mertens, U., Love, J., Lenth, R., & Christensen, R. H. B. (2021). *afex: Analysis of Factorial Experiments* (0.28-1) [Computer software]. https://CRAN.R-project.org/package=afex

Strouse, G. A., & Ganea, P. A. (2016). Are Prompts Provided by Electronic Books as Effective for Teaching Preschoolers a Biological Concept as Those Provided by Adults? *Early Education and Development*, *27*(8), 1190–1204. https://doi.org/10.1080/10409289.2016.1210457

van der Feest, S. V. H., & Johnson, E. K. (2016). Input-driven differences in toddlers' perception of a disappearing phonological contrast. *Language Acquisition*, *23*(2), 89–111. https://doi.org/10.1080/10489223.2015.1047096

van Heugten, M., & Johnson, E. K. (2014). Learning to contend with accents in infancy: Benefits of brief speaker exposure. *Journal of Experimental Psychology: General*, *143*(1), 340–350. https://doi.org/10.1037/a0032192

van Heugten, M., & Johnson, E. K. (2017). Input matters: Multi-accent language exposure affects word form recognition in infancy. *The Journal of the Acoustical Society of America*, *142*(2), EL196–EL200. https://doi.org/10.1121/1.4997604

White, K. S., & Aslin, R. N. (2011). Adaptation to novel accents by toddlers. *Developmental Science*, *14*(2), 372–384. https://doi.org/10.1111/j.1467-7687.2010.00986.x