Nordic Journal of Linguistics (2023), **46**, 406–431
doi:10.1017/S0332586522000051

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Faroese children's first words

Sissal M. Rasmussen

Námsvísindadeildin / Faculty of Education, Fróðskaparsetur Føroya / University of the Faroe Islands, Nóatún 3, FO-100 Tórshavn, Faroe Islands
Email: sissalmr@setur.fo

(Received 21 March 2019; revised 24 June 2021; accepted 9 April 2022; first published online 16 May 2022)

Abstract

Research has shown the importance of vocabulary development in relation to other parts of language development, e.g. grammar and reading development. Cross-linguistic research has shown similar as well as dissimilar tendencies regarding content in different languages. This study examines, for the first time, the characteristics of Faroese children's early productive vocabulary utilizing a Faroese adaptation of the MacArthur–Bates Communicative Development Inventories (MB-CDI). The study participants were 415 children aged 8 to 20 months. The results provide information on the composition and characteristics of lexical development in Faroese children and demonstrate that nouns are dominant among first words, as are onomatopoetic words and words describing family relationships. Faroese children are comparable to children learning other languages with respect to rate of acquisition and composition of words, with a somewhat higher share of words describing family members as stable words in the emerging language.

Keywords: Faroese; vocabulary acquisition; MB-CDI; children; infants

1. Introduction

A child's task when learning the language(s) spoken in his/her surroundings is complex and includes a vast array of different elements, such as learning the lexicon, phonological inventory, morphology, syntax, and grammatical rules of the particular language. These elements differ across languages, and individual children differ in their language learning development, although similar patterns are observed between languages and in language development (Caselli et al. 1995; Bleses et al. 2008a). Research has shown variation in vocabulary at certain age levels regarding different languages due to cross-linguistic differences, e.g. phonologically or dissimilarities regarding frequency for cultural or environmental reasons (Caselli et al. 1995; Bleses et al. 2008a). The first word is a milestone in a child's development and is influenced by a variety of aspects, e.g. the characteristics of the child, including health and cognitive factors, as well as the environment surrounding the child, such as which language is spoken to the child, whether the child is raised bilingual, and the quantity and style of talk by the child's caregivers (Hart & Risley 1995; Hoff 2006). The study of children's first words and vocabulary composition is

a growing field in linguistics, and while English is one of the most studied languages, research has also been conducted on a variety of other languages (e.g. Bleses et al. 2008a; Schults et al. 2012; Marjanovič-Umek et al. 2013; Simonsen et al. 2014).

The present study reports on the acquisition of Faroese children's first words, and data for this study were derived from a recent population-based normative study of the Faroese MacArthur–Bates Communicative Development Inventories (FAEMB-CDI; Rasmussen & Bleses 2018). Summing up the results from the norming study, which are relevant to this present study, it describes the developmental trends in early expressive vocabulary development. Regarding vocabulary size, the infants started with a mean of two words in productive vocabulary at 8 months, increasing to 120 words at 20 months. ANOVA analysis demonstrated a main effect of age ($F(12,402) = 37.2$), $p < 0.001$) for the age group ranging from 8 to 20 months. The results showed a gender difference regarding productive vocabulary in favour of girls; ANOVA analysis also demonstrated a main effect of gender ($F(1,413) = 12.5$, $p < 0.001$). In terms of numbers, girls produced a mean of 29 words over all ages in this sample, while boys produced 16 words. The study also reports that so-called *baby word forms*, which are a very common part of interactions between children and adults, are part of early vocabulary in Faroese children. Additionally, if included in the set of first words, they comprise 30% of the first 50 words, but they vanish over time (Rasmussen & Bleses 2018). The baby word forms are typically simplifications of the adult word form, e.g. *geggar* [ǧɛǧːaɪ] for *skógvar* [sǧɛǧvaɪ] 'shoes'.

1.1 Characteristics of the Faroese language and learning environment

Faroese is a language spoken by approximately 52,000 people living in the Faroe Islands (Hagstova Føroya 2019), and by an estimated additional 21,000 people outside the islands (Norðuratlantsbólkurin á Fólkatingi 2009). The Faroe Islands are situated in the North Atlantic Ocean and they are a self-governing territory within the Kingdom of Denmark. In this small language society, very limited research has been carried out regarding Faroese children's language acquisition and their early lexical development. This lack of research constrains our understanding of how children learn Faroese and our knowledge of the characteristics of their early vocabulary development. The few existing studies on language acquisition focus on older children and other aspects, such as syntax (see e.g. Heycock et al. 2013), and they are therefore not relevant to this study, which focuses on vocabulary and the characteristics of first words. The only study to date focusing on early vocabulary is the aforementioned Faroese adaptation of the MacArthur–Bates Communicative Development Inventories (FAEMB-CDI; Rasmussen & Bleses 2018). A very small study that is of interest in relation to this study, which addressed the impact that English and Danish digital input has on Faroese toddlers, showed that children in some cases spontaneously used English words when shown a picture of, for example, a flower (Steinbjørnsdóttir 2018).

Faroese is described as a West Nordic language, and compared with English, Faroese has a richer inflectional system, e.g. for nouns. Faroese has two numbers (singular and plural), three genders (masculine, feminine, and neuter), and four cases (nominative, accusative, dative, and genitive). For an in-depth description

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

of Faroese, see Thráinsson et al. (2004). One characteristic of Faroese is that there are many morpho-phonological alternations in both nouns and verbs. As an example of a noun, the word 'cow' has a low degree of transparency both morphologically and phonologically due to inflections, such as *kúgv* [kʰɪg̊v] (INDEF NOM/ACC.SG), *kýr* [kʰʊiːɹ] (INDEF NOM/ACC.PL), *kýrnar* [kʰʊitnar] (DEF NOM/ACC.PL), and *kúnni* [kʰʏnːə] (DEF DAT.SG), and these could be difficult for a child to identify as the same word. As an example of a verb, the word 'lie' undergoes changes from *liggja* [lɪdʒːa] (INF) to *lá* [lɔaː] (PST.3SG) and *ligið* [lɪːjɛ] (PST.PTCP), where it is only the onset consonant that stays the same, and the word 'walk' undergoes substantial changes from *ganga* [g̊ɛŋg̊a] (INF) to *gekk* [ʤɛʰkː] (PST.3SG) and *gingið* [g̊ɪŋʤɛ] (PST.PTCP) due to inflections in strong verbs, where one main characteristic is vowel alternation of the stem. Such morpho-phonological alternations are also present in other Germanic languages, but they are frequent and occur in central words in Faroese. For instance, see Árnason (2011) on the phonology of Faroese and, for example, Trudgill (2011), who states that:

> So not only is the amount of morphological opacity here considerable but historical linguistic studies show that it has actually *increased through time*. It is apparent that modern Faroese has in some respects become more irregular than Old Norse/Mediaeval Faroese, or indeed than Icelandic (Trudgill 2011:86–87; original emphasis).

For children acquiring the language, the segmentation task here can be challenging, and there is some indication that these alternations can affect the earliest segmentation process and temporarily slow the comprehension rate (see Rasmussen & Bleses 2018). Another characteristic of the language is a rich vowel system and a substantial difference in vocabulary use between the written and spoken language. Some spoken word forms are borrowed from Danish, while written language is more conservative regarding these word forms (Sandøy 2003; Petersen 2012). This discrepancy between oral and written language can potentially have relevance for children because children will hear different word forms when they are exposed to language based on written forms, e.g. when reading aloud from books, radio, or television programmes. Dialectal differences are found in Faroese, e.g. phonological differences in pronunciation and dialectal differences in inflectional morphology (Thráinsson et al. 2004), but Faroese is mutually intelligible for speakers. With regard to dialectal variation in Faroese and vocabulary acquisition in infants, this dialectal variation is not well researched, but it is assumed that there is not as great an impact on lexical word forms in early vocabulary acquisition. Some of the characteristics of Faroese are highlighted above, although a description of other domains that are of interest in the present study regarding child language acquisition is lacking, e.g. data on word frequency in both adult and child language and a description of syllable length in Faroese.

Regarding the language environment, the language used in interactions by children and adults is Faroese, in interactions both between children and adults, child to child, and between adults (i.e. in Faroese-speaking families); the Faroe Islands have until very recently not been a multilingual society, as 90% of the population speak only Faroese at home (Hagstova Føroya 2013). However, the size of the speech

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

community results in language learning opportunities in the mother tongue being reduced because of the limited resources available, mostly economic, to produce language-stimulating materials and programmes such as cartoons, books, children's radio, television, YouTube, and apps; therefore they receive considerable input in Danish and English through these media. Faroese children are not considered bilingual in the conventional manner, as they do not use Danish or English in everyday communication and do not receive formal instruction in other languages until the age of 9 or 10 years. However, as a result of the language environment including (and to a large extent only consisting of) Danish and English, mainly through digital media, they can be characterized as asymmetrical bilinguals if the term is defined as the ability to use two or more languages sufficiently to have limited casual conversations (Myers-Scotton 2006). There have been concerns raised, e.g. by teachers about Faroese children increasingly using English, for example, in informal communication with each other, but results regarding this phenomenon will have to be revealed by future research studies.

The Faroe Islands are a family-oriented society (Gaini 2013; Hayfield et al. 2016; Hayfield 2018; Hayfield & Schug 2019). Hayfield (2018) states 'The significance of family networks was clear in the data and the women who had lived abroad pointed out how "… there's a big difference between how much more help you have here [compared to Denmark]"' (Hayfield 2018:1148). In another study, women stated that having close support networks, such as parents, siblings, and in-laws, to help with childcare was central to the Faroese society, although they also mentioned that this is declining concurrently with grandparents working more (Hayfield et al. 2016).

In summary, Faroese can be characterized as a language that has not been extensively studied with respect to child language acquisition. As mentioned, it has a rich morphology and many morpho-phonological alternations. The language environment is quite dispersed regarding material-based input and availability of digital language input in the mother tongue due to the small language society.

1.2 Characteristics of early productive vocabulary

Vocabulary increases rapidly during the first years of childhood, and typically, children understand words before they can produce them (Fenson et al. 1994). Children usually say their first word around their first birthday, although research has indicated large individual differences (Bates et al. 1994; Fenson et al. 1994). Fenson et al. (1994) found these individual differences in vocabulary size among children to be both extensive and stable over time. Research has shown that, regarding lexical development, children seem to acquire the same words whether they are typically developing children or late talkers, although late talkers may acquire the words at a slower rate (Caselli et al. 1995). Schneider et al. (2015) found that the first words children produce are predictable from two linguistic factors: input frequency and phonological complexity of the words, while McDonough et al. (2011) found that nouns tended to dominate the first words due to the imageability of the words, which means the ease with which a word can be associated with a mental image or picture.

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

Research on children's early lexical development reveals the same tendencies in similar languages, as well as smaller differences in the first lexical word forms. Braginsky and colleagues found cross-linguistic consistency regarding the factors that predict the degree of difficulty of learning words; for example, nouns and predicates were influenced by concreteness and frequency, but function words were influenced by 'babiness' (a constructed measure of a word's association with babies), number of phonemes, and sentence length (Braginsky et al. 2019). Caselli et al. (1999) argue that children undergo four different stages in lexical development, which they identified as *routines and word games*, *reference*, *predication*, and *grammar*. The first words in children's productive vocabulary are often used to name things in the environment and are context-bound; later in development, verbs and adjectives are added, followed by function words, e.g. pronouns, prepositions, and determiners. There is a discussion on the proportion of verbs versus nouns in different languages and whether children learn a language in different ways depending on how rich the morphology is and which language is learned; for a more detailed discussion, see e.g. Gentner (1982), Caselli et al. (1995), and Tardif (1996).

Regarding the composition of productive vocabulary, social terms and nouns are found to dominate vocabulary in various linguistic environments among the first 50 to 100 words (Conboy & Thal 2006; Wehberg et al. 2007), and the use of verbs, adjectives, and function words is extremely rare until vocabularies reach 100 words (Caselli et al. 1995). Research has shown that children transition from single word utterances to two-word utterances when their vocabulary reaches approximately 50–100 words (Caselli et al. 1999). Hence the development of vocabulary is crucial to the development of grammar and has been found to predict the development of grammar better than age (Devescovi et al. 2005); in addition, lexicon size is a good predictor of later language development (Lee 2011; Bleses et al. 2016). The acquisition of early linguistic milestones is therefore of importance to explore, both in research and in clinical settings.

The phonology of children's early words can be affected by the child's ability to produce specific speech sounds and different combinations of sounds and also by the language acquired (de Boysson-Bardies 1991). Garmann and colleagues studied the acquisition of early words regarding word-initial bilabials and word length in syllables for Norwegian, Danish, Swedish, Italian, and English. They found that there was a predilection for words with initial bilabials; however, in relation to word learning the effect of word length varied depending on the language acquired (Garmann et al. 2019).

A substantial segment of research on early language development is based on parental reports, which are found to be a good and reliable method of gathering data (Bates et al. 1995; Fenson et al. 2000; Feldman et al. 2005; Law & Roy 2008). One of these parental reports is the MacArthur–Bates Communicative Development Inventories (MB-CDI; Fenson et al. 2007). Findings from MB-CDI studies have shown remarkably similar patterns in acquisition, e.g. large variability and acceleration in vocabulary during the second year (Bleses et al. 2008a). The inventories are based on parents reporting which words children are producing, but there is no information on the concrete pronunciation. However, it is possible to identify the effects of phonological factors by examining which words are or are not acquired. Garmann and colleagues used the MB-CDI to assess children's

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

phonological acquisition of word production in the aforementioned study (Wehberg et al. 2007; Garmann et al. 2019).

This study utilizes the Faroese adaptation of the MB-CDI Infant form, FAEMB-CDI I *Orð og keipur* (Words and Gestures). The parental report has also been adapted from American English to the Nordic languages, including Norwegian (Simonsen et al. 2014), Swedish (Eriksson & Berglund 1999), Danish (Bleses et al. 2008b), and Icelandic (Thordardottir & Weismer 1996). The results from the Nordic MB-CDI studies are comparable to findings from MB-CDI studies on other languages regarding vocabulary growth, individual differences, asymmetry in comprehension/production rate, and composition of early vocabulary (see Rasmussen & Bleses 2018 for a comparison of Faroese versus other Nordic languages and Bleses & Trecca 2016 for a description of three mainland Scandinavian languages).

1.3 Objectives of the current study

The motivation behind the research was to conduct the first fundamental study of early expressive vocabulary, thereby providing data that have not been obtainable previously. This information will be very helpful for further research questions and for work in developing norm-referenced tools for screening and language assessment. Based on previous research, that we see the same cross-linguistic tendencies in early acquisition but also some effect of the language and environment, the suggestion could be that close family relations and the higher degree of words changing phonologically due to inflections would affect the results. The research questions in this study relate to describing Faroese children's early vocabulary acquisition regarding word categories, age, gender, and any special characteristics for emerging word forms in Faroese and agreement with other languages: (a) Which are the top 50 expressive words first learned? (b) What are the characteristics of the first words acquired regarding phonological features, semantic categories, and lexical composition? (c) Are some words found to be stable, for example, are the same words found across different age spans in the emerging language?

2. Method

2.1 Materials

To assess Faroese children's productive vocabulary acquisition, data were collected using the FAEMB-CDI parental reports (Fenson et al. 1994; Rasmussen & Bleses 2018). This instrument is well known and has been widely used to gather data regarding children's language acquisition; the instrument makes it possible to gather large amounts of data in terms of both the number of words included and the number of participants. The questionnaire has proved to be a valid instrument to assess early language skills in children (Bates et al. 1995; Fenson et al. 2000; Feldman et al. 2005; Fenson et al. 2007; Law & Roy 2008). The method has various strengths and drawbacks. Some advantages are that the instrument does not require active cooperation from children, it is an economical instrument to use in gathering data, and the word list that is used aids parents in remembering words their child

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

might use. Disadvantages are that parents can overestimate their child's ability (Law & Roy 2008), and a higher proportion of well-educated parents often participate in the questionnaire compared with less educated parents (Fenson et al. 2000, 2007; Simonsen et al. 2014). While caution should be taken in interpreting the results, this method has proved to be very well suited to gathering data regarding infants and children up to 3 years (Law & Roy 2008).

The Faroese adaptation of the MB-CDI followed the guidelines on adaptation (see http://mb-cdi.stanford.edu), to ensure the instrument is as close and comparable as possible to the original and Nordic adaptations (e.g. the Norwegian, Danish, Swedish, and Icelandic adaptations). Some changes were made in the Faroese adaptation to reflect both the cultural and linguistic environments. In the vocabulary list, some words were replaced or added, especially in the section about animals and food, and a category called 'baby words' was added that contains words used by infants and in infant-directed speech in the Faroe Islands. For a detailed description of the Faroese version of the MB-CDI tool, the adaptation process, and the norming study, see Rasmussen & Bleses (2018).

2.2 MB-CDI I vocabulary list

In this study, the infant form of the FAEMB-CDI I (Words and Gestures) scale is used; the construction of the word list is described in Rasmussen & Bleses (2018). The data comprise the first section of FAEMB-CDI I regarding productive vocabulary, which consists of a word checklist containing 428 words divided into 21 categories, including 'food and drink' and 'toys' (see Table 1). Note that the Faroese version has an added category labelled 0, which contains baby word forms. This category is not included in the sum scores or analysed in this paper but was added as an experimental part of the norming study to provide insight into a phenomenon often observed in Faroese. Of the 428 vocabulary items on the list, 26.2% are one-syllable words while 55.3% are two-syllable words. The rest of the words in the list are three-syllable words or more (11.1% and 7.3%, respectively). Regarding phonological characteristics of the words in the list, 66.3% of the words start with a consonant, 21.4% start with a consonant cluster (85 of the words with two consonants and five with three consonants), and 12.4% start with a vowel or diphthong. Of the onset consonants in the word list, 18.3% are bilabials.

2.3 Procedure

2.3.1 Sampling

The children were identified through Landsfólkayvirlitið [The National Register], which provided contact information, e.g. addresses, names of the parents and children, and the child's birthday, and the children's parents were invited by mail to participate in the study. The data collection was part of a wider project including both the infant and toddler form of the FAEMB-CDI. As stated in Section 1.1, the population of the Faroe Islands is very limited, which affects the number of infants in the population (only approximately 600 babies are born each year). Therefore, all parents of Faroese children in the age range of 8 months to 3 years received an invitation letter with thorough information about the project and a guide on how to

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

Table 1. Categories and number of items in the vocabulary list in the MB-CDI I: Words and Gestures in the Faroese, Danish, and English MB-CDIs

MB-CDI Words and Gestures	Faroese	Danish	English
0. Baby words[a]	(27)	—	—
1. Sound effects and animal sounds	12	11	12
2. Animals (real or toy)	39	36	36
3. Vehicles (real or toy)	11	10	9
4. Toys	9	8	8
5. Food and drink	33	28	30
6. Clothing	21	21	19
7. Body parts	23	20	20
8. Small household items	40	39	36
10. Outside things[b]	13	14	27
11. Places to go[b]	16	14	—
12. People	22	30	20
13. Games and routines	20	15	19
14. Action words	53	53	55
15. Words about time	9	8	8
16. Descriptive words	37	36	37
17. Pronouns	14	11	11
18. Question words	7	6	6
19. Prepositions and locations	14	16	11
20. Quantifiers and articles	12	10	8
Total vocabulary	428 items	410 items	396 items

[a]Baby words are not counted in the total number of items, as it is an added section.
[b]'Outside things' and 'places to go' represent one category in the English MB-CDI I: Words and Gestures.

complete the questionnaire. To obtain sufficient data, there were two rounds of data collection, with ten months between them. This method resulted in a small number of children in age groups 18 to 20 months (under 15%) potentially being represented twice in the dataset but in different age groups (e.g. as an 8-month-old and 18-month-old), but this is not the case for any of the children in the age groups up to 18 months. The time lag of ten months between the rounds was chosen to address the issue that the parents' responses in the first round of data collection should not be too present in their minds during the later round.

The parents had two options for completing the FAEMB-CDI I: a password-protected website or a paper version, which the parents could request. Virtually all the parents used the website, as there was only one request for the paper version. Ninety per cent of Faroese households have Internet access (Hagstova Føroya 2013). In addition to the FAEMB-CDI parental forms, the parents also completed a

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

questionnaire with background information regarding issues relevant to language development.

To fill out the FAEMB-CDI I questionnaire, the parents were asked to mark words that their child UNDERSTOOD and words that the parents had heard the child SAY ('understood and said'). The parents were asked to mark a word even if the pronunciation deviated from the adult pronunciation (examples of accepted but deviating pronunciations were provided) and to fill out the questionnaire when the child was not present. If the parents had questions, they could contact the first author via email or telephone throughout the sampling period, and common questions from the parents were posted on the web page hosting the web form under a tab labelled 'frequently asked questions' (FAQ). There were 70 comments and questions mainly concerning technical problems, e.g. problems signing in, loss of password, and questions from parents not included in the study, e.g. if the child was older or the family was living abroad, asking if they could participate (which they could not). The system allowed parents to log in and out and to return to complete the form later, and they were able to switch between the sections of the FAEMB-CDI form until they pressed the submit button. The parents had approximately 12 days to complete the form (calculated from the day they received the invitation by letter) and a reminder was sent to the parents who had not completed the form at that point. Age was calculated from the child's date of birth to the date that the parents completed the form.

2.3.2 Response rate

The data were collected over two rounds. The response rate was 51.2% in the first round and 37.7% in the second round, which resulted in a response rate of 44.5% over the two rounds. Due to the limited cohort of children, significant effort was made to inform about the data collection through all types of media, which seems to have made an impact, as the response rate was somewhat higher than that of similar studies, e.g. the Norwegian study, which had a response rate of 37% (Simonsen et al. 2014).

2.3.3 Criteria for inclusion in the study

The children included in the present study had no documented or suspected serious developmental or health issues and had to meet the following four criteria: birth at full term (from week 36), hospital stay not exceeding four weeks, no parental concerns regarding the language development of the child, and only limited contact with languages other than Faroese. All bilingual children were excluded, which included children who had been living abroad and children who spoke another language according to their parents. Both parents had to have Faroese as their first language, or one parent had Faroese and the other had a Nordic language as their first language. Ninety per cent of the parents in the sample were monolingual. Children who had physical or mental disadvantages in acquiring language were excluded, e.g. if they had been in contact with a speech and language therapist or if the parents, kindergarten teacher, or health personnel had a suspicion of language delay, hearing problems, or serious illness. Profoundly deaf children

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

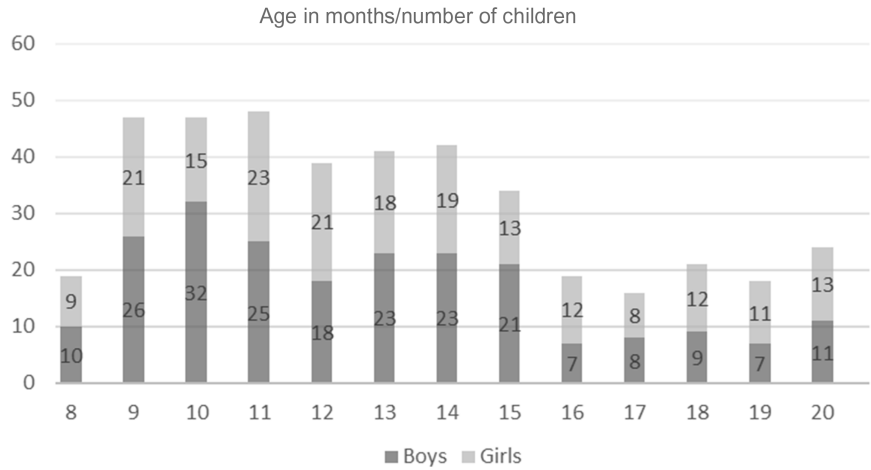

Figure 1. Number of participants in each age group.

and internationally adopted children were also excluded from the study. To be included in the study, at least one item on the questionnaire had to be marked.

2.4 Participants

The following results are based on 415 infants aged 8 to 20 months, including 220 boys (53%) and 195 girls (47%), which is in accordance with the demographic distribution of gender (Hagstova Føroya).

As shown in Figure 1, there are between 34 and 48 children in each age group from 9 to 15 months. The age groups from 16 to 20 months contain fewer children, ranging from 16 to 24, than the earlier age groups. The reason for the smaller groups from 16 to 20 months is that these data were a part of a larger study, and children in these age groups were divided across two questionnaires, FAEMB-CDI I and II, due to the research design. There are also fewer children in the 8-month age group, 19, but this may be because some parents were reluctant to answer the questionnaire when their children had only a few words or perhaps only a single word in their productive vocabulary.

2.5 Demographic characteristics of the sample

Table 2 provides information on the educational status of the parents in the final sample compared with the characteristics of the whole population.

The sample contains fewer mothers and fathers with basic education than the Faroese population in general and more with further education; this results in a sample in which parents with basic education or less are underrepresented, which is in accordance with other MB-CDI studies (Fenson et al. 2007; Bleses et al. 2008a; Simonsen et al. 2014). Regarding sibling status, a quarter of the children in this sample were the only child in the home, which is in accordance with population data, namely, 27% of families have one child (Hagstova Føroya 2014) and the rest

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

Table 2. Demographic characteristics of the sample

	FAEMB-CDI				Population	
	Father		Mother		Father	Mother
	N	%	*N*	%	%	%
Lower secondary	69	16.6	34	8.2	30.4	41.6
Vocational/high school	111	26.7	137	33.0	38.9	35.3
Medium further education	129	31.1	165	39.8	25.1	19.7
Long further education	103	24.8	79	19.0	5.6	3.4
Information missing	3	0.7	0	0		
Total	415	100	415	100	100	100

Source: Hagstova Føroya (2013).

have two or more children. The children were from all areas of the Faroe Islands, covering different dialects, and over 90% of the families were monolingual, while the remainder were bilingual, with one parent speaking a Nordic language and the other speaking Faroese. The population in the Faroe Islands is 90% monolingual (Hagstova Føroya 2013).

After applying the exclusion criteria and comparing the demographic characteristics of the final sample with the population, the sample was found to be balanced regarding monolingual status, geography, gender, and sibling status compared with the overall population but skewed toward higher educational levels of the parents.

2.6 Validity

The following results regarding the Faroese children's first words are based on an adaptation of the MacArthur–Bates Communicative Development Inventories (MB-CDI). The adaptation process, construction of the word list, and norming study are described in Rasmussen & Bleses (2018) and are summarized in the following. The reliability of the FAEMB-CDI I *Orð og keipur* (Word and Gestures) scale regarding the items in the word list was evaluated based on internal consistency scores and Cronbach's alpha, which returned a coefficient of 0.992 regarding word production for the productive vocabulary scale. This value is comparable to the results of other MB-CDI studies (e.g. Fenson et al. 2007; Bleses et al. 2008a) and demonstrates high internal consistency of the scale. Measuring the validity of the FAEMB-CDI I is difficult because of the absence of previous research on the topic or other measurement instruments in Faroese. To provide indications of the validity of the lexical items on the FAEMB-CDI, measurements from two different methods and the internal consistency for the scores on the MB-CDI I scale are provided. The first measure of validity was a composed picture test using words from the FAEMB-CDI to investigate whether frequently checked words on the MB-CDI were also produced more often by children in real testing than infrequent words (see Rasmussen & Bleses 2018 for information on the study design). The picture test contained 30 nouns, which were easy to depict, from

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

different categories, such as food, animals, and toys, in the FAEMB-CDI. Both frequent words – words that more than 75% of the children in the sample could produce – and 10 words that fewer than 25% of the children could produce were included in the picture test. The picture test was administered to 22 children who were part of the sample (parents had completed the FAEMB-CDI checklist). The results of the picture test reflected the same tendency in frequency in the picture test as in the FAEMB-CDI checklist, as the children correctly produced 8.7% of the infrequent words compared to 82.3% of the frequent words. Moreover, the FAEMB-CDI form was administered to the parents of these children, and a comparison between the picture data and the parental form showed that there is agreement between the parental answers on the FAEMB-CDI checklist and the child answers for the same words in the picture test for 70% of the items. This result confirms the agreement between the parent's responses on the FAEMB-CDI checklist and their child's actual ability to produce the words. As a supplemental measure to obtain some indication of the representativeness of words on the FAEMB-CDI, video recordings were made with two children. The parents were guided to play with books and toys and to elicit words by asking questions about, for instance, food, toys, and clothes. A comparison between the recordings and the answers on the parental form revealed that 78% of the words (different word types) said in the recordings are included in the FAEMB-CDI. To learn more about the validity and reliability of the FAEMB-CDI forms, see Rasmussen & Bleses (2018).

2.7 Scoring

Productive vocabulary is the total number of words marked by the parents as produced by the infants on the FAEMB-CDI questionnaire. The results are based on the sum scores for productive vocabulary, which is the number of words marked 'said' on the FAEMB-CDI form.

2.8 Procedure to compose the first word frequency list

In the following, the procedure to establish a list for use in describing first words will be outlined. Table 3 provides an overview of the 50 first words calculated in two different ways. Following Fenson et al. (1994), the list on the left is based on the first 50 words reported to reach the frequency of 50% in each age group. The list on the right follows Rescorla et al. (2014) in illustrating the vocabulary composition through percentage use scores and is based on the first 50 words acquired in the expressive vocabulary of Faroese children, where 'acquired' is defined as the child being reported to produce that word. The frequency of the occurrence of each word on the Infant FAEMB-CDI form was calculated, and the first 50 words were assigned a rank. These percentages are collapsed over age levels.

A comparison between the two top 50 rankings in Table 3 reveals very few differences between the two lists based on the first 50 words acquired in the expressive vocabulary of Faroese children; very few words are not the same. The differences partially stem from one list having more words than the other because all words reaching the rank of 50 (five words) are included in the list on the left, resulting in a total of more than 50 words (54 words). The list on the left has eight words

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

Table 3. Fifty first words

Rank	Word	Age in months	% of monthly sample	Rank top 50 all sample		
1	babba	10	60%	1	babba	68%
2	mamma	10	55%	2	mamma	67%
3	abbi	14	60%	3	hey	49%
4	brmm	14	52%	4	voff voff	45%
5	hey	14	52%	5	abbi	44%
6	voff voff	15	71%	6	omma	43%
7	takk	15	65%	7	brmm	42%
8	mjav	15	59%	8	takk	41%
9	omma	15	59%	9	bei	40%
10	bei	15	59%	10	mm mm	38%
11	*spæla bø*	15	56%	11	nei	38%
12	avv!	15	56%	12	mæ mæ	35%
13	mæ mæ	15	56%	13	ja	34%
14	nei	15	50%	14	*spæla bø*	34%
15	áh áh	16	53%	15	áh áh	34%
16	gvagg gvagg	16	53%	16	avv!	34%
17	skógvar	16	53%	17	mjav	33%
18	banan	17	69%	18	gvagg gvagg	30%
19	muh	17	56%	19	muh	25%
20	*navn á ped.*	17	56%	20	banan	24%
21	ja	17	56%	21	drekka (noun)	23%
22	meira	17	56%	22	skógvar	19%
23	breyð	17	50%	23	halló	19%
24	drekka (noun)	18	76%	24	tyst	19%
25	drekka (verb)	18	62%	25	breyð	18%
26	keks	18	57%	26	baby	18%
27	køka	18	57%	27	*shh/hússj*	17%
28	halló	18	57%	28	bað	17%
29	bilur	18	57%	29	heitt	17%
30	dukka	18	57%	30	meira	16%
31	eyga	18	57%	31	drekka (verb)	15%
32	baby	18	57%	32	sutta	15%
33	*sítt egna navn*	18	57%	33	*so stór/ur*	15%
34	hoppa	18	57%	34	ha?	15%

(*Continued*)

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

Table 3. (*Continued*)

Rank	Word	Age in months	% of monthly sample	Rank top 50 all sample		
35	mjólk	18	57%	35	hoppa	15%
36	bamsa	18	52%	36	bilur	15%
37	bóltur	18	52%	37	*sítt egna navn*	15%
38	bomm	18	52%	38	beiggi	14%
39	oyra	18	52%	39	eyga	13%
40	sutta	18	52%	40	klappa	13%
41	bað	19	72%	41	dukka	13%
42	ketta	19	67%	42	bomm	13%
43	hundur	19	61%	43	bóltur	13%
44	ha?	19	61%	44	hundur	13%
45	heitt	19	61%	45	keks	13%
46	jakki	19	56%	46	*navn á ped.*	12%
47	ísur	19	50%	47	bamsa	12%
48	upp	19	50%	48	mjólk	12%
49	kykkeliky	19	50%	49	ketta	12%
50	dunna	19	50%	50	bók	12%
50	bók	19	50%			
50	*shh/hússj*	19	50%			
50	eta	19	50%			
50	klappa	19	50%			

The words marked with italics represent games, e.g. *spæla bø* (game: 'peek a boo'), *sítt egna navn* (own name), which are not analysable regarding syllables and phonological characteristics, for example.

that are not included in the list on the right: *køka* 'cake', *oyra* 'ear', *jakki* 'jacket', *ísur* 'ice cream', *upp* 'up', *kykkeliky* 'cockadoodledoo', *dunna* 'duck', and *eta* 'eat' are not in the right-hand list. The list on the right has four words that are not included in the left-hand list: *beiggi* 'brother', *mm mm* 'yum yum', *tyst* 'thirsty', and *so stórur* (a game in which the child stretches his/her arms in response to 'So big?'). The results in Table 3 show that the lists are similar regarding content of words, with very few differences between them.

Fenson et al. reported high correlations between the two methods of reporting first words, and the Rescorla et al. (2014) procedure is followed in this study. Five items in the list on the right in Table 3 representing games and names marked with italics, e.g. *so stórur* 'so big' (game), *spæla bø* 'peek a boo' (game), *shh* 'shh', *sítt egna navn* (own name), and *navn á pedagogi/dagrøktara* (name of kindergarten teacher), which are not analysable regarding syllables and phonological characteristics, for example, are omitted, resulting in the final list in Table 4. This list contains 45 words (although the list is labelled 'Fifty first words', as it is based on the top 50 words), and the following results are based on this list. Regarding the baby word forms

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

Table 4. List of the 50 first words (excluding five items; see main text)

Word	Translation	Frequency	Rank all	English	Rank boys	Rank girls
babba	daddy	68.4	1	1	1	1
mamma	mummy	67.0	2	2	2	2
hey	hi	48.9	3	4	5	3
voff voff	woof woof	45.1	4	11	7	4
abbi	grandpa	44.3	5	38	4	8
omma	grandma	43.1	6	30	6	7
brmm	vroom	41.7	7	16	3	17
takk	thank you	40.7	8	32	9	5
bei	bye	40.2	9	3	10	6
mm mm	yum yum	38.3	10	13	8	14
nei	no	37.6	11	9	12	9
mæ mæ	baa baa	35.2	12	6	13	10
ja	yes	34.2	13	45	14	11
áh áh	uh oh	34.0	14	5	11	16
avv!	ouch	33.7	15	36	15	13
mjav	meow	32.8	16	25	16	12
gvagg gvagg	quack quack	29.6	17	23	17	15
muh	moo	25.3	18	31	19	18
banan	banana	23.6	19	26	18	21
drekka (noun)	drink	22.7	20		20	19
skógvar	shoes	19.3	21	24	32	20
halló	hello	19.3	22	48	21	23
tyst	thirsty	18.8	23		22	25
breyð	bread	18.3	24		26	24
baby	baby	18.1	25	12	39	22
bað	bath	17.1	26	40	27	27
heitt	hot	16.9	27	27	24	29
meira	more	16.1	28		28	31
drekka (verb)	drink	15.4	29		29	34
sutta	pacifier	15.2	30		37	28
ha?	what	15.2	31		23	64
hoppa	jump	14.9	32		34	30
bilur	car	14.7	33	44	25	45
beiggi	brother	14.5	34		38	32

(*Continued*)

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

Table 4. (*Continued*)

Word	Translation	Frequency	Rank all	English	Rank boys	Rank girls
eyga	eye	13.3	35	29	49	33
klappa	clap	13.0	36		33	47
dukka	doll	13.0	37		141	26
bomm	candy	13.0	38		31	73
bóltur	ball	12.8	39	7	42	38
hundur	dog	12.5	40	8	45	39
keks	cracker	12.5	41	37	35	74
bamsa	teddy bear	12.0	42		85	35
mjólk	milk	12.0	43		41	50
ketta	cat	11.8	44	21	71	41
bók	book	11.8	45	17	86	36

mentioned in Section 1, they could have been included in the list; however, to allow for cross-linguistic comparison and because they are not stable over time, they are not included in the list.

3. Results

3.1 First words

In this section, descriptive analyses of the results of Faroese children's first productive words are presented. As can be seen in Figure 2, the Faroese results of the productive vocabulary for infants are comparable to that of other languages. The mean number of words produced by Faroese infants is compared with that produced by children exposed to other languages via Wordbank, an open database of children's vocabulary development, using the MB-CDI questionnaire (see http://wordbank.stanford.edu/). The Swedish data include children up to the age of 16 months, and the English data include children up to 18 months, while the Faroese, Danish, and Norwegian data include infants up to 20 months. Faroese children show the same developmental trend as children exposed to the other languages and are most similar to English children and slightly higher in productive vocabulary than Norwegian children.

3.1.1 List of the first 50 words

Table 4 lists the 50 first words acquired by Faroese children (after removing five unanalysable items, as explained above). The list shows the frequency of each word on the list, a translation to English, and separate rankings for Faroese boys and girls. The rank for the word in English from the American MB-CDI study is added for comparison. The American study was chosen as it reflects a large-scale study and was the original study that other results emerge from. Blank spaces in the English

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

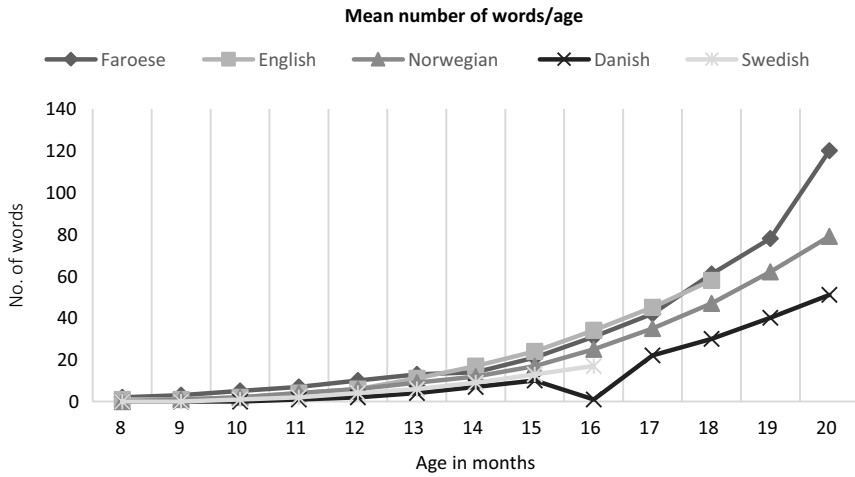

Figure 2. Productive vocabulary by language.

top 50 word list reflect that the word is found in the Faroese top 50 but is not included in the American top 50 word list.

3.1.2 Content and composition of first words

The next section is an analysis of the first words produced by the children in this study regarding content and composition, and it was determined whether any differences can be seen in ranking of words regarding gender. The first words were also compared with English children's first words.

Word types/categories. The FAEMB-CDI I word list is divided into 20 different categories, and eight of the categories are represented in Table 4. As for the lexical categories the list has 23 nouns (51%), which include some onomatopoetic words and words used in social routines, and three verbs (7%), two adjectives (4%), and the remainder of the words are onomatopoetic expressions, interjections, and adverbs. The three verbs Faroese infants use in their productive vocabulary are *klappa* 'clap', *drekka* 'drink', and *hoppa* 'jump'.

It is not possible to use the MB-CDI to assess how children produce words phonologically, although it is possible to analyse the words that are acquired in terms of their phonological characteristics. The results in Table 4 show that the word list contains 30 words starting with a consonant, with 20 of these words (44%) starting with a bilabial and five words (11%) starting with a vowel (one of these is a diphthong). Eight of the words, equal to 16%, start with a consonant cluster. Regarding syllable length, the list contains 27 disyllabic words (60%, including onomatopoetic words such as *mm mm*) and 18 monosyllabic words (40%).

The distribution of words from different semantic categories reveals a tendency for words to belong to three categories: people, sound effects and animal sounds, and games and routines. This finding is illustrated in Table 5 in selected age groups,

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

Table 5. Ten most frequent words and their distribution across categories in productive vocabulary

	10 months	15 months	20 months
Sound effects and animal sounds	3	5	4
People	3	4	3
Games and routines	5	3	4

as only words from these three categories are present among the 10 most frequent words (the numbers do not sum to 10 because some words share the 10th rank).

First words and gender. Analysing the word for gender differences reveals that the highest-ranked words in the list are quite similar; in the top 20 words there is only one word that is not the same for both genders (the girls have the word 'shoe' in the top 20 instead of 'banana', although *banan* is ranked as 21, so it is near the top 20). Regarding differences between the genders in ranking in the top 20 words, the word *brmm* 'vroom' has the most different rankings: it is ranked 3 for boys and 17 for girls.

From the top 50 list for both genders combined, there are five words that are not included in the boys' list: *eyga* 'eye', *dukka* 'doll', *bamsa* 'teddy bear', *ketta* 'cat', and *bók* 'book'. There are also five words that are not on the girls' list but are on the list of the top 50 words for both genders: *ha?* 'what?', *klappa* 'clap', *bomm* 'candy', *keks* 'cracker', and *mjólk* 'milk'. Thus, 40 of the 45 words on the list of first words are the same for both genders, resulting in 88.9% of the 50 first words being the same for both genders.

First words by language. Comparing the Faroese and English lists of the first 50 words, we see that 69% of the words are the same, with the difference between ranks in Faroese and English being 10 ranks or less for 40% of the words. Among the 10 words with rankings that are more than 20 places apart in the two languages, two of these words are *omma* 'grandma' and *abbi* 'grandpa', which have notably different ranks, e.g. numbers 5 and 6 in Faroese compared with 38 and 30 in English.

3.1.3 Emerging word forms

Table 6 displays the emerging word forms based on the top 10 list of words for each of the months represented in this study. The percentage reported is for individual words produced by the children month by month from 8 months to 20 months. For the 8-month age group, the list represents ALL different words said at that point, but for the remainder of the age groups, the top 10 words are shown. In some age groups, more than 10 words are shown, which is due to more than one word having the same rank in the list and therefore sharing a rank number.

Regarding the emerging productive vocabulary in Table 6, these words are predominantly from the 'people' category, which includes family members, and the 'games and routines' category. In 10 of the 13 age groups, four of the words in the top 10 represent family members. Words from the 'games and routines' category are well represented, as are 'sound effects and animal sounds'.

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

Table 6. Emerging word forms by month

8		9		10		11		12		13		14		15		16		17		18		19		20	
1. mamma	47%	1. babba	40%	1. babba	60%	1. babba	54%	1. mamma	74%	1. mamma	68%	1. mamma	67%	1. babba	85%	1.mamma	89%	1. mamma	88%	1. babba	95%	1. babba	94%	1. voff voff	100%
2. babba	42%	2. mamma	36%	2. mamma	55%	1. mamma	54%	2. babba	72%	2. babba	66%	1. babba	67%	2. mamma	79%	2. voff	74%	2. voff voff	81%	2. hey	95%	1. mamma	94%	1. babba	100%
3. abbi	21%	3. spæla bø	19%	3. hey	38%	3. mm mm	40%	3. mm mm	49%	3. brmm	49%	3. abbi	60%	3. voff voff	71%	2. omma	74%	2. babba	81%	3. avv!	90%	1. hey	94%	1. takk	100%
4. systir	11%	3. hey	19%	4. mm mm	26%	4. brmm	35%	4. hey	46%	4. voff voff	46%	4. brmm	52%	4. takk	65%	4. brmm	68%	2. omma	81%	3. voff voff	90%	1. takk	94%	4. mæ mæ	96%
4. brmm	11%	5. mm mm	17%	5. voff voff	19%	5. hey	31%	5. spæla bø	41%	5. hey	44%	4. hey	52%	5. abbi	62%	4. abbi	68%	5. bei	75%	5. omma	86%	5. bei	89%	5. mjav	92%
4. omma	11%	5. bei	17%	5. nei	19%	6. abbi	27%	5. brmm	41%	6. takk	41%	6. voff voff	45%	5. mjav	59%	6. avv!	63%	5. mjav	75%	5. mamma	86%	5. mjav	89%	5. avv!	92%
4. hey	11%	7. abbi	11%	7. ja	17%	7. áh áh	25%	5. omma	41%	6. omma	41%	6. omma	45%	5. omma	59%	6. mæ mæ	63%	7. banan	69%	5. ja	86%	5. voff voff	89%	5. omma	92%
8. áh áh	5%	7. omma	11%	7. spæla bø	17%	7. voff voff	25%	8. abbi	38%	6. spæla bø	41%	6. takk	45%	5. hey	59%	6. bei	63%	7. abbi	69%	7. mjav	81%	5. omma	89%	5. bei	92%
8. mammubeiggi	5%	7. so stór	11%	7. abbi	17%	7. omma	25%	8. voff voff	38%	9. mæ mæ	39%	9. bei	43%	8. bei	59%	6. hey	63%	7. takk	69%	7. banan	81%	5. nei	89%	5. hey	92%
8. hatta	5%	10. beiggi	9%	10 brmm	15%	10. takk	23%	8. bei	33%	9. abbi	39%	10. nei	40%	9. spæla bø	56%	6. takk	63%	10. avv!	63%	7. abbi	81%	10. mæ mæ	83%	5. mamma	92%
		10. brmm	9%	10. bei	15%					9. nei	39%			9. avv!	56%			10. áh áh	63%	7. bei	81%	10. ja	83%	5. nei	92%
		10. nei	9%											9. mæ mæ	56%			10. gvagg gvagg	63%	7. nei	81%	10. abbi	83%		
																		10. hey	63%						

Table 7. Emerging word forms produced at different ages

		Percentages of children producing word forms by age (months)				
Word	Translation	16	14	12	10	8
babba	daddy	**89%**	67%	72%	60%	42%
mamma	mummy	**89%**	67%	74%	55%	47%
voff voff	woof woof	**74%**	45%	38%	19%	0%
omma	grandmother	**74%**	45%	41%	11%	11%
brmm	vroom	**68%**	52%	41%	15%	11%
abbi	grandfather	**68%**	60%	38%	17%	21%
avv!	ouch	**63%**	31%	23%	6%	0%
mæ mæ	baa baa	**63%**	38%	28%	2%	0%
bei	bye	**63%**	43%	33%	15%	0%
hey	hello	**63%**	52%	46%	38%	11%
takk	thank you	**63%**	45%	28%	13%	0%
nei	no	**58%**	40%	26%	19%	0%
áh áh	uh oh	**53%**	36%	23%	13%	5%
gvagg gvagg	quack quack	**53%**	24%	21%	9%	0%
skógvar	shoes	**53%**	14%	5%	0%	0%

Analysing the top 10 list according to age reveals that some words are stable over the age spectrum, e.g. the six words *mamma* 'mummy', *babba* 'daddy', *bei* 'bye bye', *omma* 'grandmother', *abbi* 'grandfather', and *brmm* 'vroom' are present in six of the 13 age groups. Five of these six words are present in other age groups. When comparing the first age group, 8 months, with the last, 20 months, four of the top 10 words are the same. This result demonstrates a very stable vocabulary based on the most frequent words from 8 months to 20 months. Other emerging word forms, such as *bei* 'bye bye', can also be traced through the table and are present in 10 of 13 age categories. Therefore they are quite stable among the most common words.

In the case of the words in the 8-month category, there are two words to comment on. The word *systir* 'sister' is in the top 10, and because of its phonologically complex form, having a word-medial consonant cluster, an explanation is required. In the 8-month group there are only 19 children. Few children say anything at this time, and, as reported previously, the list contains all words said at this point. Thus, if the word is checked by the parents, it ends up in the top 10 list, which makes the data more unreliable in the 8-month-old group. Another explanation is that parents actually mean that the child says the easier word form *didda*, which is the baby word form for *systir*. Both forms have the same frequency, perhaps because parents 'translate/replace' the word *systir* with *didda*. Regarding the word *mammubeiggi* 'uncle', which is a phonologically complex

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

four-syllable word, either the case is the same as with *systir* in that the child says *mamba* (baby word form for *mammubeiggi*), or the data are more unstable, for example, if only two parents check the word, it ends up in the top 10 because there are few children saying anything at this time.

Table 7 shows emerging word forms from the FAEMB-CDI I, focusing on the age group of 16 months, which is the oldest age group for which the form was originally intended. The total sum score for the words produced by more than 50% of those children is selected. Then the percentage for the same words for selected ages (14, 12, 10, and 8 months) is shown to trace how the word behaves in younger age groups, showing the developmental trend of the word in relation to age.

All of the words show a steady increase with age. The words *mamma* 'mummy' and *babba* 'daddy' are both represented with the highest percentages for all age groups, and there is a tendency for the ranking of words at 16 months to be quite similar to the rankings in the younger age groups.

4. Discussion

This large-scale study presented findings on the first words in Faroese children's early lexical development. The main goal was to describe the composition of early productive vocabulary, which to date is unexplored with respect to frequency, emerging words, syllable length, age, gender, and use of different categories, using the Faroese adaptation of the MB-CDI. As there is very restricted research in the Faroese language regarding both adult- and child-based data, the comparison is based on the characteristics of the MB-CDI I word list and cross-linguistic research.

The trend observed in other languages, namely that verbs, adjectives, and grammatical function words are rare in early lexical acquisition, was also observed in this study (Caselli et al. 1995, 1999). The results show that, as for other languages, nouns (including kinship terms, onomatopoetic words, and common nouns) are dominant among Faroese children's first words, with over 50% of the first words being nouns and 7% being verbs. This could be due to input frequency, e.g. found by Schneider et al. (2015) as stated in Section 1, ease of articulation, and imageability of the word, but these factors have not been measured in Faroese. Onomatopoetic words and names for persons are considered as nouns, e.g. *mamma* 'mummy', *abbi* 'grandfather', and *gvagg gvagg* 'quack quack', and they have the highest frequencies while common nouns dominate the lower frequencies. The three verbs among the first 50 words also reflect the finding in international research that the first verbs are often verbs reflecting activity, in this case 'clap', 'drink', and 'hop'. Regarding phonological characteristics of the words the children are reported to produce, we see similarities between the words in the total list of 428 words and the first 50 words. Eleven per cent of the first 50 words start with a vowel or diphthong, compared with 12.4% in the full word list; 60% of the words start with a consonant compared with 66.3% in the full word list; and 16% start with a cluster of two or three consonants (85 of the words with two consonants and five with three consonants), compared with 21.4% in the full word list. These numbers are very close, with a somewhat smaller proportion of consonant clusters than that for the whole set. As revealed in Table 4,

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

Faroese infants have a preference for words starting with initial bilabials, as nearly half of the first words they produce start with a bilabial, while 18.3% of the words in the word list are bilabials. Regarding syllable length, the children produce 27 disyllabic words (60%, including onomatopoetic words such as *mm mm*) and 18 monosyllabic words (40%), compared with 55.3% disyllabic words and 26.3% monosyllabic words in the total word list (the rest of the words in the list are three-syllable words or more, 11.1% and 7.3%, respectively). The infants seem to learn both monosyllabic and disyllabic words, with a larger proportion of disyllabic words, in the first stages of productive vocabulary, as seen in international research, e.g. Danish MB-CDI studies (Bleses et al. 2008b). If the results are compared to the study by Garmann et al. (2019), which comprises data on children's early words regarding both initial bilabials and syllable length in Norwegian, Swedish, English, and Italian, the Faroese children follow the tendency to use word-initial bilabials, but regarding syllable length they resemble the Swedish and Italian children. Although word frequency lists for adult or child language are not found in Faroese, the results of the first 50 words demonstrate no surprise or unexpected items, but words that contain the predicted characteristics of frequency, social routines, and non-complex phonological characteristics, e.g. words for family members such as *mum* and *dad*, words for animal sounds such as *moo* and *woof woof*, and nouns such as *car*, *doll*, *book*, and *banana*. The first 50 words are mostly not words influenced by morpho-phonological alternations described as a characteristic of Faroese. Only three of the first 50 words are characterized as such words: *skógvar* [sgɛgvaɹ] 'shoes', which in an inflectional form *skóm* (DAT.PL) is pronounced [sgɔun], *drekka* [ɖɹɛʰkːa] 'to drink' as [ɖɹʊʤɔː] (*drukkið*, PST.PTCP), and *bók* [bɔuk] 'book' as [bøːkɹ] (*bøkur*, NOM/ACC.PL). It would be expected that the infants learn the easier and more frequent word forms first compared with more opaque word forms due to these morpho-phonological alternations.

Comparing the 50 first words with the English-American children's first words shows that 69% of the words are the same, which reflects a tendency that the first words are universal as stated in previous research. Analysis of the list regarding words that have quite different rankings reveals that words for family members, such as 'grandpa' and 'grandma', have quite different rankings, e.g. 5 and 6 in Faroese compared with 30 and 38 in English, as well as *beiggi* 'brother', which is one of the first 50 words in Faroese but not in English. This could reflect input frequency stemming from a more family-oriented society, as stated in Section 1 (Hayfield et al. 2016; Hayfield 2018), although these words are easier to pronounce in Faroese compared with English, e.g. they are less complex phonologically and comprise early-acquired segments such as the bilabials /m/ and /b/. This result is also in accordance with the Wehberg et al. (2007) findings regarding the later acquisition of the words for 'mother' (Danish *mor*) and 'father' (Danish *far*) compared with the English words *mum* and *dad*. Other words with different rankings in the two languages include *takk* 'thank you', which in Faroese is ranked 8, compared with 32 in English. This could stem from the circumstance that in English, it is phonologically more complex and is a two-word utterance and therefore more complicated than the Faroese word *takk*, although children in the sample are able to utter two-syllable words, e.g. *grandma* (in this case two-word utterances are not overly more complicated than a two-syllable word). However, the words *ball* and

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

dog in English are ranked 7 and 8, as opposed to 39 and 40 in Faroese, where they are two-syllable words, *bóltur* 'ball' and *hundur* 'dog', and phonologically more complicated to produce. As stated in Section 1, the first words children produce can be predicted from linguistic factors such as concreteness, input frequency, babiness, phonological complexity, and imageability of the words, and these results show that the words acquired by Faroese children early are seemingly also related to these factors, although it is not possible to relate them to research on word frequency and babiness in Faroese, for example.

The results provide the first information on the characteristics of vocabulary acquisition in Faroese children and add to the literature by providing descriptive data. In summary, we found a domination of nouns and social terms among the first 50 words. In terms of emerging word forms, the highest ranked words are quite stable across age groups, indicating that these words are important words in early vocabulary. Words reflecting family are present throughout all ages and may reflect a family-oriented environment, as seen in Italian, where children have a higher proportion of social terms, e.g. terms for family members, than English children in the earliest stages of language acquisition. Caselli and colleagues explained the higher representation of social terms with the tendency for Italian children to have family closer, and, therefore, both have a reason for naming them and prompting more interaction (e.g. Caselli et al. 1995, 1999; Hayfield et al. 2016). At this stage, children do not use many verbs and follow the trend seen in other languages that first verb forms reflect actions, e.g. clapping and jumping. We can see some small gender differences in the composition and ranking of vocabulary, although not in the 20 highest ranked words, which seem to be the same for both genders. These results are comparable to other Scandinavian results and to English results, and may be an expected result given that the cultures are fairly similar and considering the method used, namely, the word lists are adaptations of the same instrument. Rescorla et al. (2014) showed that some words in the early word learning process have strong cross-linguistic similarities and the present study demonstrates that they are also part of the early Faroese vocabulary.

It must be taken into account that the demographic data show that the sample is skewed regarding the educational status of the parents in relation to the population, and education is known to have an impact on children's language development. The validity of the assessment method has been proved, but it is based on parents and can be influenced by memory limitations. The sample contains a higher proportion of parents with education compared to the population. Parental education influences children's language acquisition and therefore this element must be considered when interpreting the results. It should be noted, as stated in the description of the participants, that the sample included 44.5% of the whole population in this age spectrum, although the sample is smaller in number than those in Norwegian and Danish studies.

There are some limitations to this study, mainly, the lack of other research to validate the data and the reliance of the results on parental reports rather than direct measures of vocabulary acquisition. Therefore generalizations of the results are limited to parent-reported vocabulary. Despite these limitations, the study has important value for constructing language tests, for example, and identifying future directions that research on Faroese child language should take.

https://doi.org/10.1017/S0332586522000051 Published online by Cambridge University Press

The next research steps should involve analysing broader aspects of lexical acquisition regarding age, and include receptive vocabulary, longitudinal data, and other aspects of language acquisition.

Acknowledgements. The author would like to thank all the parents participating in this study, as well as Philip Dale and the CDI Advisory Board for providing permission to adapt the MacArthur–Bates Communicative Development Inventories to Faroese. I want to thank the three anonymous *NJL* reviewers and editor Gunnar Ólafur Hansson for valuable feedback and comments.
The author has no conflicts of interest to declare.
This research was funded by a grant from Granskingarráðið (The Faroese Research Council).

References

Árnason, Kristján. 2011. *The Phonology of Icelandic and Faroese*. Oxford: Oxford University Press.

Bates, Elizabeth, Philip S. Dale & Donna Thal. 1995. Individual differences and their implications for theories of language development. In Paul Fletcher & Brian MacWhinney (eds.), *The Handbook of Child Language*, 96–151. Oxford: Blackwell.

Bates, Elizabeth, Virginia Marchman, Donna Thal, Larry Fenson, Philip Dale & J. Steven Reznick. 1994. Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language* **21**(1), 85–123.

Bleses, Dorthe, Guido Makransky, Philip S. Dale, Anders Højen & Burcak Aktürk Ari. 2016. Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics* **37**(6), 1461–1476.

Bleses, Dorthe & Fabio Trecca. 2016. Early acquisition of Danish in a cross-Scandinavian perspective: A psycholinguistic challenge? In Hans Olav Enger, Monika I. Norvik Knoph, Kristian E. Kristoffersen & Marianne Lind (eds.), *Helt fabelaktig! Festskrift til Hanne Gram Simonsen på 70-årsdagen* [Absolutely fabulous! Festschrift for Hanne Gram Simonsen on her 70th birthday], 13–28. Oslo: Novus forlag.

Bleses, Dorthe, Werner Vach, Malene Slott, Sonja Wehberg, Pia Thomsen, Thomas O. Madsen & Hans Basbøll. 2008a. Early vocabulary development in Danish and other languages: A CDI-based comparison. *Journal of Child Language* **35**(3), 619–650.

Bleses, Dorthe, Werner Vach, Malene Slott, Sonja Wehberg, Pia Thomsen, Thomas O. Madsen & Hans Basbøll. 2008b. The Danish Communicative Developmental Inventories: Validity and main developmental trends. *Journal of Child Language* **35**(3), 651–669.

de Boysson-Bardies, Bénédicte. 1991. Adaptation to language: Evidence from babbling and first words in four languages. *Language* **67**(2), 297–319.

Braginsky, Mika, Daniel Yurovsky, Virginia A. Marchman & Michael C. Frank. 2019. Consistency and variability in children's word learning across languages. *Open Mind* **3**, 52–67.

Caselli, Maria C., Elizabeth Bates & Paola Casadio. 1999. A comparison of the transition from first words to grammar in English and Italian. *Journal of Child Language* **26**(1), 69–111.

Caselli, Maria C., Elizabeth Bates, Paola Casadio, Judi Fenson, Larry Fenson, Lisa Sanderl & Just Weir. 1995. A cross-linguistic study of early lexical development. *Cognitive Development* **10**(2), 159–199.

Conboy, Barbara T. & Donna J. Thal. 2006. Ties between the lexicon and grammar: Cross-sectional and longitudinal studies of bilingual toddlers. *Child Development* **77**(3), 712–735.

Devescovi, Antonella, Maria Cristina Caselli, Daniela Marchione, Patrizio Pasqualetti, Judy Reilly & Elizabeth Bates. 2005. A crosslinguistic study of the relationship between grammar and lexical development. *Journal of Child Language* **32**(4), 759–786.

Eriksson, Mårten & Eva Berglund. 1999. Swedish early communicative development inventories: Words and gestures. *First Language* **19**(55), 55–90.

Feldman, Heidi M., Philip S. Dale, Thomas F. Campbell, D. Kathleen Colborn, Marcia Kurs-Lasky, Howard E. Rockette & Jack L. Paradise. 2005. Concurrent and predictive validity of parent reports of child language at ages 2 and 3 years. *Child Development* **76**(4), 856–868.

Fenson, Larry, Elizabeth Bates, Philip Dale, Judith Goodman, J. Steven Reznick & Donna Thal. 2000. Reply: Measuring variability in early child language: Don't shoot the messenger. *Child Development* **71**(2), 323–328.

Fenson, Larry, Elizabeth Bates, Philip S. Dale, Virginia A. Marchman, J. Steven Reznick & Donna J. Thal. 2007. *The MacArthur–Bates Communicative Development Inventories: User's Guide and Technical Manual*. Baltimore, MD: Paul H. Brookes Publishing.

Fenson, Larry, Philip S. Dale, J. Steven Reznick, Elizabeth Bates, Donna J. Thal & Stephen J. Pethick. 1994. Variability in early communicative development. *Monographs of the Society for Research in Child Development* **59**(5), i–185.

Gaini, Firouz. 2013. *Lessons of Islands: Place and Identity in the Faroe Islands*. Tórshavn: Faroe University Press.

Garmann, Nina G., Pernille Hansen, Hanne G. Simonsen & Kristian E. Kristoffersen. 2019. The phonology of children's early words: Trends, individual variation and parents' accommodation in child-directed speech. *Frontiers in Communication* **4**, 10.

Gentner, Dedre. 1982. Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In Stan A. Kuczaj (ed.), *Language Development*, vol. 2: Language, Thought and Culture, 301–334. Hillsdale, NJ: Lawrence Erlbaum.

Hagstova Føroya [Statistics Faroe Islands]. Retrieved 25.10.19.

Hagstova Føroya. 2013. [Statistics Faroe Islands] *Øll telja við, 2013* [Everybody is accounted for].

Hart, Betty & Todd R. Risley. 1995. *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore, MD: Paul H. Brookes Publishing.

Hayfield, Erika A. 2018. Family-centred work motility in a small island society: The case of the Faroe Islands. *Gender, Place & Culture* **25**(8), 1138–1153.

Hayfield, Erika A., Rógvi Olavson & Lív Patursson. 2016. *Part-time Work in the Nordic Region III: An Introductory Study of the Faroe Islands, Greenland and Åland Islands*. Copenhagen: Nordic Council of Ministers.

Hayfield, Erika A. & Mariah Schug. 2019. 'It's like they have a cognitive map of relations': Feeling strange in a small island community. *Journal of Intercultural Studies* **40**(4), 383–398.

Heycock, Caroline, Antonella Sorace, Zakaris S. Hansen & Frances Wilson. 2013. Acquisition in variation (and vice versa): V-to-T in Faroese children. *Language Acquisition* **20**(1), 5–22.

Hoff, Erika. 2006. How social contexts support and shape language development. *Developmental Review* **26**(1), 55–88.

Law, James & Penny Roy. 2008. Parental report of infant language skills: A review of the development and application of the Communicative Development Inventories. *Child and Adolescent Mental Health* **13**(4), 198–206.

Lee, Joanne. 2011. Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics* **32**(1), 69–92.

Marjanovič-Umek, Ljubica, Urška Fekonja-Peklaj & Anja Podlesek. 2013. Characteristics of early vocabulary and grammar development in Slovenian-speaking infants and toddlers: A CDI-adaptation study *Journal of Child Language* **40**(4), 779–798.

The MacArthur–Bates Communicative Development Inventories. http://mb-cdi.stanford.edu

McDonough, Colleen, Lulu Song, Kathy Hirsh-Pasek, Roberta M. Golinkoff & Robert Lannon. 2011. An image is worth a thousand words: Why nouns tend to dominate verbs in early word learning. *Developmental Science* **14**(2), 181–189.

Myers-Scotton, Carol. 2006. *Multiple Voices: An Introduction to Bilingualism*. Oxford: Blackwell Publishing.

Norðuratlantsbólkurin á Fólkatingi. 2009. *Eg vil so fegin hava, at børnini læra føroyskt* [I would so like for the children to learn Faroese]. https://old.setur.fo/fileadmin/user_upload/SSS/PDFfilur/Strategi/Fragreidingar/Utisetakanningin_i_Dk_3partur.pdf

Petersen, Hjalmar P. 2012. *The Dynamics of Faroese–Danish Language Contact*. Heidelberg: Universitätsverlag Winter.

Rasmussen, Sissal M. & Dorthe Bleses. 2018. Faroese children's early vocabulary acquisition: A Faroese adaptation of the MacArthur–Bates Communicative Development Inventories. *First Language* **38**(6), 641–668.

Rescorla, Leslie, Alessandra Frigerio, Maria E. Sali, Pietro Spataro & Emiddia Longobardi. 2014. Typical and delayed lexical development in Italian. *Journal of Speech, Language, and Hearing Research* **57**(5), 1792–1803.

Sandøy, Helge. 2003. Purisme i færøysk [Purism in Faroese]. In Helge Sandøy, Randi Brodersen & Endre Brunstad (eds.), *Purt og reint: Om purisme i dei nordiske språka* [Pure and clean: On purism in the Nordic languages], 85–109. Volda: Høgskulen i Volda.

Schneider, Rose, Daniel Yurovsky & Mike Frank. 2015. Large-scale investigations of variability in children's first words. In David C. Noelle, Rick Dale, Anne S. Warlaumont, Jeff Yoshimi, Teenie Matlock, Carolyn D. Jennings & Paul P. Maglio (eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, 2110–2115. Austin, TX: Cognitive Science Society.

Schults, Astra, Tiia Tulviste & Kenn Konstabel. 2012. Early vocabulary and gestures in Estonian children. *Journal of Child Language* **39**(3), 664–686.

Simonsen, Hanne G., Kristian E. Kristoffersen, Dorthe Bleses, Sonja Wehberg & Rune N. Jørgensen. 2014. The Norwegian Communicative Development Inventories: Reliability, main developmental trends and gender differences. *First Language* **34**(1), 3–23.

Steinbjørnsdóttir, Eyð B. 2018. *'Lítla mín hvat sigur tú? Purple, flower, blóma, blue.' Ein kanning av málinum hjá tvey–fýra ára gomlum føroyskum børnum – við serligum atliti at útlendskum orðum/íkasti.* ['Little child what do you say? Purple, flower, blóma (flower), blue.' An investigation of language in 2–4 year old Faroese children – with special attention to foreign words/input]. MA thesis, University of the Faroe Islands.

Tardif, Twila. 1996. Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology* **32**(3), 492–504.

Thordardottir, Elin T. & Susan E. Weismer. 1996. Language assessment via parent report: Development of a screening instrument for Icelandic children. *First Language* **16**(48), 265–285.

Thráinsson, Höskuldur, Hjalmar P. Petersen, Jógvan L. Jacobsen & Zakaris S. Hansen. 2004. *Faroese: An Overview and Reference Grammar*. Tórshavn: Føroya Fróðskaparfelag.

Trudgill, Peter. 2011. *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: Oxford University Press.

Wehberg, Sonja, Werner Vach, Dorthe Bleses, Pia Thomsen, Thomas O. Madsen & Hans Basbøll. 2007. Danish children's first words: Analysing longitudinal data based on monthly CDI parental reports. *First Language* **27**(4), 361–383.

Wordbank. http://wordbank.stanford.edu.

Cite this article: Rasmussen SM (2023). Faroese children's first words. *Nordic Journal of Linguistics* **46**, 406–431. https://doi.org/10.1017/S0332586522000051