The role of production in infant word learning

Marilyn May Vihman
University of York

Studies of phonological development that combine speech-processing experiments with observation and analysis of production remain rare, despite the fact that production experience is necessarily relevant to developmental advance. This talk will focus on three levels of word learning, at each of which production appears to affect ‘intake’ from the speech stream.

(1) Articulatory filter: The idea that infants ‘select’ first words to say at least partially on the basis of how pronounceable they are derives from the surprising finding that a child’s very first words are relatively accurate. The construct of an ‘articulatory filter’, designed to account for this finding, is supported by several recent experiments.

(2) U-shaped curve: Once a child has established a small expressive lexicon she is able to generalize the production patterns she uses most (whether this is understood as a purely sensorimotor ‘procedure’ or a more abstract process of secondary distributional learning), adapting less accessible target words to her emergent template(s); this results in a regression in accuracy but supports an advance in word learning.

(3) Lexical knowledge facilitates new learning. An effect of expressive vocabulary on new word learning is alluded to in passing in several studies, despite the fact that correlations of word recognition with vocabulary size are rarely seen. The effect is transitory: At given points in lexical development (not at set ages) a child’s productive knowledge of words supports memory for new word forms. I will provide some evidence for this claim, which is congruent with the idea that knowledge is gradient, involving an increase in stability and reliability with repeated exposure and use.
The role of production in perception of mispronounced words

Lena Renner
Stockholm University

First language acquisition could be investigated from two different domains: perception and production. Earlier perceptual studies showed different reactions to correct word forms compared to mispronounced ones, indicating that lexical representations of early word forms are stored in detail (e.g. White & Morgan 2008, Swingley 2003). However, these perceptual experiments do not take the individual production of the child into account. My research project involves different studies on perception and production. At the Norphlex-seminar in Tromsø I will present the results of an eye-tracking study on the perception of mispronounced words and I hope to be able to present some results on the ongoing study of production.
An analysis of mothers' production of onomatopoeia in IDS
Catherine Laing
University of York

Onomatopoeic forms are consistently found to be prolific in infants’ first words, but despite their prominence these forms remain largely unaccounted for in child language research, and are often disregarded from analyses of child language development on this basis. Kunnari (2002) suggested that onomatopoeia may be prosodically more salient than non-onomatopoeia in the input, but does not follow up this claim with any further discussion. This study seeks to determine why these forms are so frequent in the early output on the basis of Kunnari’s proposal, through an acoustic analysis comparing the prosodic features of onomatopoeic and non-onomatopoeic words in infant-directed speech (IDS). Here it is hypothesised that onomatopoeia will be produced with more salient prosodic features than non-onomatopoeic forms.

Onomatopoeic words (OWs, e.g. quack) and their corresponding conventional words (CWs e.g. duck) were extracted from recordings of 12 British mothers reading picture books to their 8-month-old infants. Stimuli were measured in Praat and OW-CW pairs were compared for features typically reported in IDS: mean pitch, pitch range, word duration, pause duration and word-repetition (Fernald & Kuhl, 1987).

OWs were found to be significantly more salient than their CW equivalents across all typical features of IDS, supporting Kunnari’s (2002) suggestion that onomatopoeia may be prolific in infant speech due to their acoustic salience in IDS. This highlights the important relationship between perception and production in early language development, and further demonstrates the effect of IDS features on infant language development.

Validation of MacArthur-Bates CDI-III:
A pilot study of monolingual and bilingual children between 3 and 4
Janne von Koss Torkildsen
University of Oslo

The MacArthur-Bates Communicative Development Inventories (CDIs) are parental report forms assessing language and communication skills in children between 8 and 36 months. These reports have been validated and normed in a variety of languages, and are widely used as screening tools. However, there is a need for parental reports which can provide adequate measures of the language abilities of preschool children older than 36 months. A CDI form for children between 36 and 48 months has been developed and adapted to some languages (e.g. American English, Swedish, Danish and Norwegian), but there are no published validation studies of this form.

The present pilot study aimed to validate the Norwegian adaptation of a new MacArthur-Bates Communicative Development Inventory (CDI-III) developed for children between 36 and 48 months. The validation was performed by comparing the scores on a web based version of the CDI III with scores on behavioral measures of language abilities in a small sample of monolingual and bilingual children. The behavioral data was obtained from a combination of standardized tests (British Picture Vocabulary Scale II and auditory sequential memory from Illinois Test of Psycholinguistic Abilities) and measures calculated from recordings of spontaneous speech (mean length of utterance, lexical variation, grammatical errors).

Results showed that there were significant correlations between CDI subtests measuring grammatical skills and linguistic complexity on the one hand and corresponding behavioral measures on the other. On the CDI vocabulary measure, however, there was a tendency towards a floor effect, and the expected correlations with behavioral tests were not found. The findings in this pilot study indicate that further adjustments to the Norwegian CDI-III are required before a full scale validation can be conducted.
The Swedish Early Communicative Development Test is a succession of the Swedish Early Communicative Development Inventory and is a parent report form designed for normal children aged 30 to 48 month. It assesses children’s vocabulary, syntax and orthographic skills. The development of the form is described. Reports from 1300 children have been attained. The psychometric data for the included scales was good. Age (in months) explained a considerable part of the variation in each scale. A preliminary validation against the Children’s Communicative Checklist-2 yielded satisfactory results.
Adapting the MacArthur-Bates' CDI to North Sámi - minority language issues

Berit Anne Bals Baal & Kristine Bentzen
Institutt for språkvitenskap/CASTL, UiT - Norges arktiske universitet

One of the main objectives of the DASAGO project ([http://site.uit.no/dasago/nb/](http://site.uit.no/dasago/nb/)) is to adapt the MacArthur Bates Communicative Development Inventory (CDI henceforth) to North Sámi. North Sámi is the largest of the Sámi languages, and it is spoken in Norway, Sweden and Finland. The number of speakers is estimated to be around 15000-25000, and most of these live in Norway. The CDI is a language assessment tool based on parent reporting. It is used for assessing the early lexical growth and linguistic development in children, and it consists of two forms. Part of the CDI project of DASAGO is to develop norms of typical language development in North Sámi, and for that subproject we need native speaker participants. In addition to the linguistic challenges in developing the CDI (Baal & Bentzen 2013), there are also some challenges connected to the minority language status of the North Sámi language. In this presentation we focus on the low number of native speakers, bilingualism among the children acquiring the language, the impact of the majority language on the language use of the parents and the on-going language preservation measures of the North Sámi communities. We will look at the kind of questions these issues raise, particularly the choice of language both for user instructions and lexical items and the selection of participants later on in the project.

References
Developing a CDI for Urdu and Pashto

Mariam Dar
University of York

The MacArthur Communicative Development Inventory (CDI) is a standard checklist on which parents record their children’s lexical progress. CDIs have been developed for American, British English, French, Catalan, Italian, Japanese, Spanish and Russian, among other languages, but not for Urdu or any other regional language in Pakistan. However, the multilingual situation in Pakistan makes development of a CDI for Urdu and Pashto a laborious undertaking as children in typical Pakistani households are exposed to two or more languages. Urdu and English co-exist in almost every middle-class household. Urdu is the national language and English is the official language of the country and also the medium of instruction in schools. This explains the common occurrence of English loan words in Urdu and other regional languages like Pashto. Children are exposed to cartoons and children’s books in English even before going to school. Apart from English and Urdu, at least one regional language is spoken in many homes (sometimes two, if the mother and father have different regional languages).

Another postgraduate and I started off by interviewing mothers from Urdu- and Pashto-speaking households in Pakistan to develop a comprehensive list of words spoken and understood by 1.5 – 3 year olds. British and American CDIs were also taken as a foundation for the set of categories and the final lists were sent to several mothers from Urdu- and Pashto-speaking homes for feedback and corrections. The Urdu and Pashto CDIs that we formulated consequently not only include English and regional equivalents of Urdu words related to everyday life but also are very much reflective of life in Pakistan, with words related to topics such as religion, traditional foods, clothing, expressions of approval/disapproval, places to visit – and power cut-offs, a regular experience in Pakistani life. The development of Pashto and Urdu CDIs will pave the way for future research, making it convenient to collect data for those working on lexical development in these languages. It should also be of interest to practitioners working in the areas of education and speech therapy, not only in Pakistan but also in places with large numbers of Pakistani emigrant families.
Prosodic-phonetic biases and their role in the acquisition process

Elinor Payne
Oxford University

When learning to speak, infants’ exposure to the particular distributional frequencies (in e.g. contrastive consonant length, syllable structure, etc.) of the ambient language can subtly influence the developmental path followed (Vihman & Velleman, 2000; Prieto et al, 2006) despite common neuro-physical constraints, as the infant capitalizes on that language’s statistical properties to learn its structures (Saffran et al, 2003). However, languages also exhibit interesting micro-variation in their implementation of phonological structure, patterns that I refer to as prosodic-phonetic biases (PPBs).

In this paper I discuss the potential role of PPBs in shaping the language acquisition process, and describe a project in progress* that examines the emergence of PBBs in Norwegian and English and focuses in particular on the co-ordination and timing of consonants. Preliminary findings from English are reported.

*The Acquisition of Consonant Timing: Payne, Post, Simonsen, Garmann & Moen; funded by the British Academy and the University of Oslo
Vihman and Croft (2007) argue that many children use phonological templates when they acquire language. Templates are phonological patterns, and when a child uses templates, his or her word productions become more similar to each other than to the target words. The templates are linked to the child’s babbling patterns and/or their first words, and are based on the sounds that the child already knows how to produce. The child uses his or her templates to select words that are easy for him or her to acquire, and to adapt words to make them easier to remember, articulate and plan.

Iben (1;5) may serve as an illustration. She produces bisyllabic words that have a word initial bilabial. She says pappa ‘daddy’, baby ‘baby’ and pippip ‘child word for bird’ in a target-like way, but when she says eple ‘apple’ or søpla ‘the garbage’, she adapts her productions to fit her template, saying [bela] and [pøleje], respectively. Iben also shows a lot of variation in her productions. During a half hour video recording, she says the word eple six times, every time in a different way. Even so, all her productions fit her bisyllabic word-initial bilabial template (Garmann 2012).

To analyse templates, one can use the transcript of a 30-minute recording of the child playing with one or both parents. Then, one picks one exemplar of each word type, or more if the child uses different prosodic shapes for the same word, and groups the word types according to prosodic shapes. If at least 10 per cent of the word types in the recording have a similar phonological shape, the prosodic shape is considered to be a template.

Another argument for a specific template would be if the template is shown to be stable over some time. In our talk, we will analyse the templates in four consecutive monthly recordings of four Norwegian monolingual children to see whether they have templates that they use over time.

Vihman and Croft (2007) claim that templates are individual by child, and at the same time that they are influenced by the children’s input. The study will therefore also say something about what is typical for Norwegian children’s templates.

Literature


Prosodic focus marking in Swedish speaking children

Anna Sara Romøren
Utrecht University

According to the Lund model of Stockholm Swedish intonation, focus is marked by adding a floating high tone to the lexical accent of a focal word, and by downstepping post-focal accents.\(^1\) Research on prosodic accent development in Swedish children has primarily been concerned with word prosody,\(^2\) thus little is known about the acquisition of sentence prosody. To shed light on this issue we used a picture game to investigate prosodic focus marking in Swedish children, compared to adults.

We elicited SVO sentences from 10 Stockholm Swedish-speaking 4-to-5 year-olds (mean age 5;0) and 10 adult controls (mean age 26;7) through an interactive picture-game. Thirty question-answer mini-dialogues were embedded in the game to elicit sentences with narrow focus in initial, medial, and final position, in addition to contrastive focus medially and broad focus on the whole sentence. For the analysis, the presence of focal high tones was manually coded, and the coding was automatically extracted using Praat.\(^3\) In addition, measurements of maximum pitch, minimum pitch and word duration were extracted, combining automatic and manual measures.

Mixed effect modelling was used to examine the effect of focus (focal vs. post-focal) on the presence of a focal H in medial and final position. Both groups were found to consistently mark focus by adding a focal H, but the children occasionally also added this tone to non-focal targets, particularly in medial position. Our phonetic analysis revealed both groups to expand the pitch range for focus in medial position, but only the adults also did this sentence finally. Both groups increased the duration on focal target words, but sentence-medially this effect was stronger in the adults than in the children.

Our results suggest that Swedish children are well on their way to adult proficiency in terms of adding the focus marking high to focal constituents. However, the use of pitch range and duration for marking focus differed between the children and the adults, suggesting that the development toward adult proficiency continues beyond the age of 4- to 5. We will discuss our results in light of comparable data from Dutch.\(^4\)

---

4 Romøren A.S.H., & Chen A. Accentuation, pitch and duration as cues to focus in Dutch 4-to 5-year-olds. 2014;38.
Multilingualism represents a common language development path in Europe due to enhanced migration. Bilingual children are prone to have smaller vocabularies in at least one language compared to monolinguals (Bialystok, Luk, Peets, & Yang, 2010), which means that typical bilingual language development can be hard to disentangle from Specific Language Impairment (SLI) (Leonard, 2013). This study investigates lexical skills in Polish-Norwegian preschoolers, compared to their monolingual peers, using the Cross-Linguistic Lexical Tasks (CLT), a new assessment tool constructed within the European program COST IS0804 to be comparable across a wide range of languages.

The CLT assess comprehension and production of nouns and verbs through picture choice and naming. The tasks, currently ready for 20 different languages, are based on morphological and phonological complexity and subjective age of acquisition (AoA) for a set of words with meanings shared across 34 languages. Culturally-neutral colored pictures were made specifically for the CLT.

We have investigated lexical skills in 27 bilingual children living in Norway with at least one Polish parent, 27 monolingual Norwegians and 65 monolingual Polish children, all aged between 3;0 and 6;11. Parents have filled in a background questionnaire about their child’s development and language input.

After a brief presentation of the construction of the CLT instrument, we will present results from bilinguals and monolinguals, focusing on the effects of psycholinguistic properties, language use in the family and SES on their lexical skills.

References:


Assessing exposure effects in gender assignment by Norwegian-Russian bilingual children

Yulia Rodina & Marit Westergaard
MultiLing/University of Oslo, CASTL/University of Tromsø

This study is a follow-up to Rodina & Westergaard (2013) and investigates the acquisition of grammatical gender by simultaneous Norwegian/Russian bilingual children with the special focus on two factors: transparency of the gender system and the amount of parental input. Both Norwegian and Russian distinguish masculine, feminine and neuter gender, but differ in transparency of gender assignment: While Russian gender is to a large extent predictable based on a few morphophonological rules, Norwegian gender is relatively opaque. The focus of this study is on adjective-noun agreement in Russian, illustrated in (1), and gender marking on indefinite articles in Norwegian, illustrated in (2).

(1) a. zelenyj dom
green.M house(M)
‘a green house’
b. zelenaja ruka
green.F arm(F)
‘a green arm’
c. zelenoje okno
green.N window(N)
‘a green window’

(2) a. en grønn kopp
a.M green cup.M
‘a green cup’
b. ei grønn såpe
a.F green soap.F
‘green soap’
c. et grønt hus
a.N green house.N
‘a green house’

The frequency of exposure which is found to play a crucial role in bilingual gender acquisition (cf. Gathercole & Thomas 2005, Unsworth et al. 2012) was measured in two ways. First, the children in our study were from two types of households: families with two Russian-speaking parents and families with one Norwegian- and one Russian-speaking parent. The amount of exposure was also measured by using the newly developed tool called UBiLEC, Utrecht Bilingual Language Exposure Calculator (Unsworth 2013).

We have conducted an elicited production experiment with 20 Norwegian/Russian children (4;1-7;11) and compared their developmental patterns to those of 20 monolingual Norwegian (4;4-6;0) and 20 monolingual Russian children (4;2-6;0). The results of the study show that the amount of exposure in terms of parental input has a significant effect on the bilingual children’s agreement accuracy in Russian. Bilingual children with one Russian-speaking parent perform significantly worse than the children with two-Russian speaking parents and monolinguals (p=0.001). Our results also show that Norwegian-Russian children with one Russian-speaking parent are also different qualitatively, as not only opaque but also transparent nouns are error-prone in their production. To explain these results we have considered the role of parental input patterns as well as two other input measures obtained through the UBiLEC.

References:
Gradual acquisition of grammatical function morphemes has long been a topic in child language studies. Yet the pure-syntactic explanations often fail to consistently account for this gradual learning process. The parameter-setting approach predicts that the acquisition of a grammatical structure occurs instantaneously, and studies couched within that approach have proposed that sporadically appearing functional morphemes are either ‘optional’ (e.g. Wexler 1994) or ‘imposters’ (i.e. parts of rote-learned constructions, not properly established in child’s grammar, cf. Radford 1990, 1996). Recent studies in phonological acquisition suggest that this variability may at least partly be due to phonological constraints, and could be accounted for in terms of developing prosodic representations in child language (see Demuth 2001, 2007, Gehrken 1994, 1996 among others).

In this talk, I explore the possibility that the omission of prepositions at the telegraphic speech stage may in fact be an artifact of a specific ranking of phonological constraints. In lines with recent research on L1 acquisition within Optimality Theory (e.g. Gnanadesikan 2004, Fikkert 2007), I argue that children start with a ranking where Markedness constraints override Faithfulness constraints. On the next stage, Positional Faithfulness constraints are promoted above General Markedness constraints, which is reflected in the production of unstressed elements in phonologically prominent positions (immediately pretonic syllables in Russian). The ranking of Markedness constraints above General Faithfulness constraints is responsible for the deletion of unstressed elements in other positions. I provide a statistical analysis of elicited production data from a group of monolingual Russian children (n=34; aged 2;4–3;9 years), which shows a significant difference in the probability of preposition omission depending on the position of the preposition within a prosodic word. Young children who omitted prepositions in their speech were more likely to do so when the preposition was followed by an unstressed syllable than when it occupied a first pretonic position.

References
CDI used in language screening of mono- and bilingual children in Denmark

Dorthe Bleses & Anders Højen

University of Southern Denmark

According to a review by Law et al. (2008) and a overview of CDI adaptations to different languages (Dale & Penfold, 2005), one important application of CDI’s is in screening for language delay or disorders. In Denmark, a short list version of the CDI has been used in a national language screening program of mono- and bilingual children. Here we will address the strengths and limitations of such an approach, with a particular focus on the use of CDI in bilinguals.

Is Danish difficult to segment?

Dorthe Bleses, Anders Højen (Fabio Trecca, Hans Basbøll, Thomas O Madsen, University of Southern Denmark & Morten Christiansen, University of Southern Denmark and Cornell University)

Previous research has shown that Danish children’s development of comprehension skills is relatively slow compared to a range of other languages. Low segmentability of Danish, due to its phonetic structure rich in vocoids and long vocalic stretches, has been suggested as a possible cause (e.g., Grønnum, 2003; Gooskens et al. 2010). As a point of departure for a later discussion of cross-language differences in segmentation or acquisition difficulty, we present preliminary data from three different sources: 1) an segmentation experiment where adult subjects from four different language backgrounds where tested on their ability to recognize words in an artificial “contoid” or “vocoid” language; 2) distribution of diphone transitional probabilities in Danish and American speech input to children; 3) computer modelling of segmentation of Danish and English based on the distributional characteristics of Danish.
This presentation will discuss some preliminary results from an EEG study where potential changes in the cortical activity associated with the early lexical development were investigated longitudinally in a group of Swedish infants. The initial overall age range was 1-24 months but consistent data could not be obtained for the complete age range. The current results are based on 21 infants who were 425 days old (SD=43.7, range 395 to 490 days) on the first session, participated in the study. Forty high frequency nouns were selected from the typical receptive SECDI test battery. For each word, an audio recording and a drawing of the object referred by the word were created. The drawings and the words were presented to the infants in both lexically congruent and lexically incongruent pairs. The infant’s EEG activity was measured in a 2×2 factorial test design – congruency and the infant’s receptive knowledge of the word (“known”/“unknown” status specified by the infant’s caregiver before each test session). There is a significant interaction between known and unknown lexical items in congruent and incongruent audio-visual stimuli presentations but this seems to be limited to monosyllabic words. Possible explanations for why monosyllabic words would be treated differently than polysyllabic words are being considered. The possible role of acoustic or semantic characteristics of the monosyllabic words in relation to the polysyllabic is being considered.
In this paper we present some new data on gender agreement in a Norwegian dialect (Tromsø), collected from several groups of speakers: adults, teenagers, and three different age groups of children. Two experimental studies have been carried out, testing the use of indefinite articles as well as double definite forms, illustrated in (1). Experiment 1 investigates all three genders, while Experiment 2 focuses on four subgroups of feminine nouns. Findings from Experiment 1 show that there is massive overgeneralization of the masculine indefinite article (en) in the child data, argued to be due to the lack of transparency of gender assignment. The neuter gender seems to be acquired (at 90% accuracy) around the age of 7, making gender a late acquisition in Norwegian compared to other languages. The feminine gender, on the other hand, is hardly used at all in the children’s production, while the adults use it 100% and the teenagers around 60-70%. At the same time, declensional endings (represented by the suffixal definite article) are acquired early in all three genders. Experiment 2 shows a similar picture with respect to differences between the age groups. Furthermore, the production of feminine gender forms is relatively low in the child data, for all four subclasses of feminines, with a slightly (but statistically significant) higher proportion of feminine forms used with nouns with female reference.

We interpret these findings as an ongoing change in progress, involving loss of the feminine indefinite article, possibly also feminine gender altogether, affecting subclasses of feminine nouns at slightly different rates. The change is presumably due to sociolinguistic factors, but we argue that the nature of the change is due to the process of language acquisition, the relevant factors being extensive syncretism (between masculine and feminine), lack of frequency, lack of transparency, as well as early acquisition of declensional forms compared to agreement. The result of this process is a simplification in the gender system, which is accompanied by added complexity in the declensional system, in that the new common gender has two different declensional classes, as illustrated in (2).

(1)  
\[\text{ei rød flaske – den røde flaska}\]
\[\text{a red bottle – the red bottle}\]

(2)  
\[\text{en bil – bilen}\]
\[\text{a.COMM car – car.DEF}\]
\[\text{‘A car - the car.’}\]

\[\text{en bok – boka}\]
\[\text{a.COMM book – book.DEF}\]
\[\text{‘A book – the book.’}\]