Estimates of Infants' Vocabulary Composition and the Role of Adult-instructions for Early Word-learning

Klintfors E., Lacerda F. & Sundberg U.
Department of Linguistics, Stockholm University

Abstract

The purpose of this study was to assess characteristics of the growth of language in Swedish 7-15 months-old infants. More scpecifically, to estimate vocabulary composition (the 1st part of the study), information on 24 infants' comprehension and production of speech was collected, and to investigate the role of adult-instructions for early word-learning (the 2nd part of the study) video recordings of infant-adult interaction-dyads were produced. The vocabulary-data were collected based on parental reports using the Swedish version of the MacArthur Communicative Development Inventory. The subjects' size of receptive and productive vocabulary across five semantic categories were estimated in a longitudinal perspective based on 45 completed forms. The infant-adult interaction data was likewise collected in a longitudinal perspective. The variables chosen for analyses of infant-adult interaction were: objective of the parent's gestural instruction, objective of the parent's vocal instruction, and focus of attention of the infant. The 1st part of the study showed an immense progress for comprehension within the receptive semantic categories of 'phrases' and 'sounds'. The results for the 2nd part of the study showed a shift along a time axis from gestural instructions being mostly correlated with the attentional focus of the infant (at 7 months) to gestural and vocal instructions being equally common (at 15 months). A discussion on significance of these results will round off the paper.

Introduction

This study is performed within two interdisciplinary projects: "Modeling Interactive Language Learning"[1] (MILLE, supported by the Bank of Sweden Tercentenary Foundation) and "Learning and Development of Contextual Action"[2] (CONTACT, supported by the European Commission). The human subject part of the projects uses behavioural data from experiments on infants' speech perception, production and adult-infant interaction (SU), from studies on infants' motoric development (UU), as well as from parallel investigations at the neuronal level (UNIFE). The non-human animal part of MILLE uses data from speech perception studies on mammals' capacity of generalization (CMU). And finally, mathematical models (KTH), and humanoid-robotic systems (UNIGE, IST) are constructed to simulate infants' performances.

The aim of the current study is twofold: to explore how adults conceive their infant's linguistic competence and to investigare the role of adult-instructions for early word-learning. For these purposes a questionnaire study was conducted and video recordings of adult-infant interaction-sessions were produced. The results are expected to be of interest for implementation of speech development in computational (MILLE) and embodied systems (CONTACT).

[1] Dept. of Linguistics, Stockholm Univ. (SU, Sweden), Dept. of Psychology, Carnegie Mellon Univ. (CMU, USA), and Dept. of Speech, Music and Hearing, Royal Inst. of Technology (KTH, Sweden).

[2] Dept. of Psychology, Uppsala Univ. (UU, Sweden), Dept. of Linguistics, Stockholm Univ. (SU, Sweden), Dept. of Communication, Computer and System Sciences and Laboratory for Integrated Advanced Robotics, Univ. of Genova (UNIGE, Italy), Neurolab and Dept. of Biomed. Sciences – Human Physiology, Univ. of Ferrara (UNIFE, Italy), Computer Vision Lab and Dept. of Electrical & Computer Engineering, Inst. Superior Técnico of Lisbon (IST, Portugal).

Background

In the view of the proposed Ecological Theory of Language Acquisition (ETLA) young children attach labels to items on the basis of general-purpose perceptual and memory mechanisms (Lacerda & Sundberg, 2006). The infant is assumed to receive multi-sensory (visual, auditory, olfactory, gustatory, tactile etc.) information available in the environment and to temporarily store simultaneous incoming information received through the different sensory channels. In a wider time-prespective, repeated co-occurences of eg. word-object pairings will permanently be stored, while possible initial wrong semantical assumptions are erased in the absence of repetition. This kind of multi-sensory associative processes – from which further linguistic referential function may emerge – are futher argued to be able to be observed even in non-human species.

On the one extreme, language acquisition may be alleged to be determined by the innate capacities of an individual (Chomsky, 1975; Pinker, 1994). As opposed to this view ETLA, just like other interactionist-approaches (Bates et al., 1998; Karmiloff-Smith, 1996), accentuate the interaction of genes and environment to determine the outcome of language.

What biologically determined propensities are there to perceive or generate sounds in particular ways? As an example, due to the short pharyngeal tract in relation to the oral cavity observed in the newborn infant (Fort & Manfredi, 1998), the spectral characteristics of the infant's utterances (in addition to be proportionally shifted to higher frequencies) are necessarily not proportioned to the adult's. Thus, despite of analogous anatomic and physiologic characteristics of articulatory gestures, the acoustic outputs of the infant and adult are unlinearly related.

In line with the social-pragmatic view on language acquisition highlighting the role of parent-child interaction – we assume that an important factor for how children break the language barrier is that the language learner interacts with his/her ecological environment. The above exemplified biological biase thus most likely determines the initial attempts of the infant to babble, but these vocalizations are perceived and influenced in line with the expectations of the caregiver. Adults change their speaking strategies in response to their assumptions on how to meet the infant's communication needs: after reaching the initial objective of maintaining communication (cf. Jakobson, 1968 phatic function of speech communication), adults often use prosodic and lexical repetitions, expanded intonation contours, longer pauses between utterances, and even exaggerate pronunciation at the segmental level (van der Weijer, 1999; Kuhl et al., 1997; Sundberg & Lacerda, 1999).

Method

The subjects visited the babylab of SU approximately once/month. Each visit started off with a perception experiment (eye-tracking-method), and then after a video-recording of adult-infant interaction, the experimenter filled the questionnaire based on parental information.

Vocabulary Composition

The subjects were 24 Swedish infants (13 girls, 11 boys, age range 6.3-15.1 months, mean 10.7 months) randomly selected from the National Swedish address register (SPAR) on the basis of age and geographical criteria. All the subjects were primarily exposed to Swedish in their families. The subjects were not paid to participate in the study.

Data on subjects' vocabulary composition was collected using The Swedish Early Communicative Development Inventory (SECDI) (Eriksson & Berglund, 1996). The inventory is a Swedish version of the MacArthur Communicative Development Inventory (CDI) (Fenson, 1993) using parental reports. The "words and gestures" form of the inventory, which is primarily designed for children 8-16 months of age, was used. The parent (or experimenter) marked "understand" and/or "understands and says" for the relevant items for each section of the SECDI-inventory. All affirmative answers per section were added to obtain the infant's comprehension and production scores. Occasionally (e.g. if the infant showed signs of fatigue) the parent filled the form at home, and sent it back to the laboratory within about a week.

The analyses was based on 45 completed forms divided in three age groups as follows: 15 (7-9 month-olds, mean 7.67), 18 (10-12 month-olds, mean 10.94), and 12 (13-15 month-olds, mean 14.10). The ratio of infants assessed repeatedly at different ages was 24:45 (6 completed one form, 12 two forms, and 5 three forms) resulting in a mix of longitudinal and

cross-sectional data. The questionnaire-data was analyzed across the following five semantic categories:

- 'Phrases' (eg. *Look over there*)
- 'Sounds' (eg. *bä bä*)
- 'People' (eg. *mother*)
- 'Small items' (eg. *a watch*)
- 'Time expressions' (eg. *later*).

Adult-infant Interaction

Data from one infant and her parent was video recorded on three occasions (à 1 hour). On these occasions the infant was 7, 11, and 15 months. Five minutes/recording (selected five minutes after the onset of each recording to assure that the play-session had properly started) were labelled in the annotation software Anvil (4.5) with reference to the following variables:

- Objective of the parent's gestural instruction (what the parent is showing)
- Objective of the parent's vocal instruction (what the parent is saying)
- Focus of attention of the infant (what the infant is attending/looking at)

Results

Vocabulary Composition

The results for receptive vocabulary showed a furthermost progress for the semantic category 'Phrases': the subjects comprehension of phrases (such as *Are you hungry?*, *Be careful*, or *Could you go and get…* etc.) increased from about 10% at the youngest age (7-9 months) to about 60% at the oldest age (13-15 months). Similar progresses can be observed for the categories 'Sounds' (such as *kvack kvack*, *nam nam*, *tuff tuff* etc.) and 'People' (such as *a baby*, *an boy*, *a lady* etc.). The semantic categories 'Small items (such as *keys*, *a tootbrush*, *a flower* etc.) and 'Time expressions' (such as *tonight*, *now*, *tomorrow* etc.) showed the least progress – only a couple of subjects at the oldest age (13-15 months) comprehended a few of these words.

The greatest progress for productive vocabulary composition occurred within the semantic category 'Sounds'. As compard with the subjects *comprehension* of 'Sounds' the *production* of expressions belonging to this category was more modest. Some progress can also be observed for the semantic categories 'Small items' and 'People' at the oldest age (13-15 months). The category showing no productive progress was 'Time expressions'.

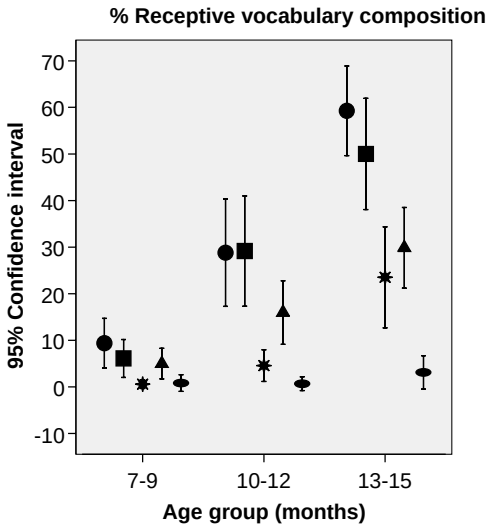

Figure 1. Receptive vocabulary composition (7-15 months). The five comprehension-categories are plotted as follows: 'Phrases' (circles), 'Sounds' (squares), 'People' (triangles), 'Small items' (stars), and 'Time expressions' (ellipses).

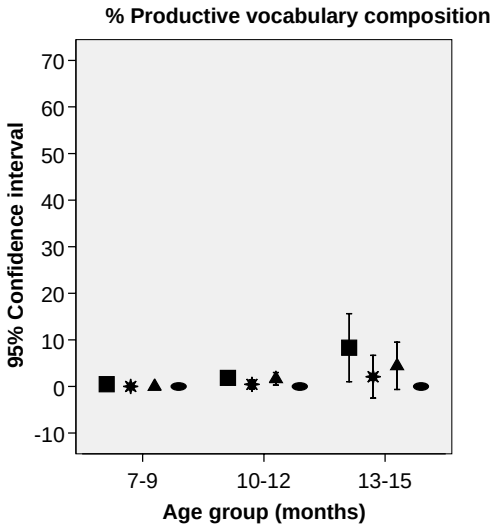

Figure 2. Productive vocabulary composition. The four production-categories are plotted using the same symbols as in Figure 1. The category 'Phrases' was measured for comprehension only and is thus absent in this figure.

Adult-infant Interaction

An analysis of correlation between Objective of the parent's gestural instruction (what the parent is showing), Objective of the parent's vocal instruction (what the parent is saying) and Focus of attention of the infant (what the the infant is attending/looking at) showed that at the youngest age (7 months) the infant mostly

directed her gaze towards what the parent was showing. By 11 months of age the infant attended more to what the parent was saying. And finally, when the infant was 15 months she attended equally to what the parent was showing and/or saying.

Table 1. Type of adult instruction mostly correlated with Focus of attention of the infant.

Age	Type of adult-instruction
7	Gestural instruction
11	Vocal instruction
15	Gestural and Vocal instruction

Discussion

In this study we have focused on the beginnings of word learning at 7-15 months. The theoretical framework, recognizing the importance of genes-environment interaction for early word learning, is provided by the Ecological theory of language acquisition. Within ETLA, early multisensory-associative processes guided by biologically determined perception/production propensities, as well as the role of infant-adult interaction are emphasized. Thus, as opposed to a single speech specific mechanism, parallel general-purpose perceptual/memory procesesses are held responsible for lexical development.

At approximately 19 months of age childrens' vocabulary rapidly expands. But the aim of this study was instead to investigate vocabulary composition at the outset of mastering words. Interestingly, indicating that word learning may start with recognition of embedded target-words, the greatest receptive progress was observed for the category of 'Phrases'. The SECDI inventory used allows in total for estimation across 19 semantic categories. In future, scores for each of these (as opposed to the five, respectively four categories analyzed in the current study) might obviously be computed to study more detailed vocabulary composition. For studies of relations between major parts of speech in child language a division in common nouns, predicates and close class words is also achievable.

If the conundrum of Quine (Quine, 1960) is to be taken as relevant to the problem of word learning, the results from the 2nd part of the current study are easily interpreted. The problem of infinite number of word-world mappings at 7 months is alleviated by gestural instructions of the adult. At 15 months, due to a narrowed search space of the infant, gestural and vocal instructions may be equally used. In addition to study how infants' focus of attention changes depending on adult-instructions (gestural or vocal), a future query of importance would be to investigate the role of adults' attunement to their perception of infant's intent.

Acknowledgements

Work supported by the Swedish Research Council, the Bank of Sweden Tercentenary Foundation, and the European Commission. The authors are greatful to a group of students in Linguistics for initial preparation of data and analysis of the video-materials.

References

Bates E, Elman J, Johnson M, Karmiloff-Smith A, Parisi D, & Plunkett, K (1998). Innateness and Emergentism. In Bechtel W & G Graham (eds), *A Companion to Cognitive Science* , 590-601. Oxford: Basil Blackwell.

Chomsky N (1975). *Reflections on language*. New York: Pantheon.

Eriksson M & Berglund E (1996). Swedish Early Communicative Development Inventory -- words and gestures. *First Language, 19,* 55-90.

Fenson L (1993). MacArthur Communicative Development Inventories: User's Guide and Technical Manual. Ref Type: Generic

Fort A & Manfredi C (1998). Acoustic analysis of newborn infant cry signals. *Med.Eng Phys., 20,* 432-442.

Jakobson R (1968). *Child language. Aphasia and phonological universals.* (72 ed.) The Hague: Mouton.

Karmiloff-Smith A (1996). *Beyond Modularity - A Developmental Perspective on Cognitive Science.*

Kuhl P, Andruski J E , Chistovich I A , Chistovich L A, Kozhevnikova E V, Ryskina V L et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277,* 684-686.

Lacerda F & Sundberg U (2006). An Ecological Theory of Language Acquisition. *Linguística, 1,* 53-106.

Pinker S (1994). *The Language Instinct: How the Mind Creates Language*. (1 ed.) New York: William Morrow and Company, Inc.

Quine W V O (1960). *Word and object*. Cambridge, UK: Cambridge University Press.

Sundberg U & Lacerda F (1999). Voice onset time in speech to infants & adults. *Phonetica, 56,* 186-199.

van der Weijer J (1999). *Language Input for Word Discovery.* MPI Series in Psycholinguistics.