Word Frequency Influences Sound-Meaning-Associations

Eeva Koponen
Department of Linguistics, Stockholm University

Abstract

The infant's ability to establish correspondences between target-words in an artificial language and their referent objects was examined in Swedish five to fifteen month-olds. The 'Visual Preference' paradigm was used in three experimental conditions implemented by video films showing objects along with matched presentation of sentences of different degrees of complexity: one highly variable speech sample and two low-variability speech samples matched to counterbalanced video presentations. The results showed significant between-group differences suggesting that the target-words in the low variability nonsense phrases were successfully associated with the referenced objects.

Background

The focus of language acquisition studies has recently shifted towards early learning processes where the current debate addresses two types of learning mechanisms: Altmann, Dienes & Gao (1995); Saffran, Aslin & Newport (1996) believe in statistical learning mechanisms, in which word segmentation is performed by *counting* or by mechanisms which act like Neural Network models, whereas Marcus, Vijayan, Bandi Rao & Vishton (1999) defend that a fundamental task of language acquisition is to extract and generalize *algebraic rules*.

For instance Saffran *et al.* (1996) habituated eight month-old infants with continuos speech stream (*bidakupadotigolabu..*) made of randomly ordered three-syllable nonsense words, where the high (1.0) transitional probabilities between syllable pairs within nonsense words (e.g. *bida*), were contrasted with low (0.33) transitional probabilities for between word syllables (e.g. *kupa*). One half of the test stimuli were consistent with the habituation items (e.g. *bidaku*), and the other half were made of the syllables heard during habituation but in different order (e.g. *dabiku*) or joined together by the final syllable of a word and the first two syllables of another word (e.g. *kupado*). The results of *Saffran et al*. showed longer listening times for words that were inconsistent with the habituation items.

In order to show that counting would not suffice to learn the rule that was generating sequences of words, Marcus *et al.* (1999) habituated seven month-old infants to sequences according to ABA (e.g. *ga ti ga*) or ABB (e.g. *ga ti ti*) 'grammars'. Half the test items were constructed from the same 'grammar' as the one the infant was habituated to, and half the test trials were constructed from the 'grammar' on which the infant was not trained. The test items were made of entirely new words, such as w*o fe wo* or *wo fe fe*. With respect to the habituation items, these new test sentences transitional probabilities were zero. The results of *Marcus et al*. showed longer listening times for words generated according to the 'grammar' on which the infants were not trained.

The purpose of the present study was to go beyond the acoustic signal per se – to find out if infants manage to extract words from continuous speech streams associating novel target-words to arbitrary objects. The hypothesis was that exposure to a nonsense language, characterized by the typical 'repetitive' structure (high target-word frequency rate, target-words in phrase final position) of infant-directed speech (IDS), would be helpful in establishing correlation between signal (speech sounds) and another co-occurring (visual) sensory dimension.

The nonsense words used as test material were randomly generated according to context-free word rules based on a description of phonotactics and morphotactics of Swedish IDS. The nonsense words corresponding to different parts-of-speech (POS) and phrase structural positions were put together in declarative, interrogative, and imperative phrases according to context-free phrase rules based on the syntax of Swedish IDS. In contrast to synthetically produced syllables (*Saffran et al*., 1996; *Marcus et al*., 1999; *Gomez & Gerken*, 1997) the nonsense phrases were read aloud by a human speaker. The *naturally* produced phrases were aimed to reflect infants sensitiveness to linguistic stimuli.

The assumption was that if we learn that the grammatical object comes after the verb from simple sentences like *Har du sett nallen?* (*Have you seen the teddybear?*), then we can generalize that regularity to grammatical objects that are very different in sound or meaning (**Fur bu skrett kullan?*), or in length (**Fur bu skrett böm blina, föna kullan?*). In line with this assumption it is interesting if statistical inferences can be used to acquire grammar. The hypothesis was that just like phonotactic constraints, also morphotactic constraints, POS and syntax form statistical regularities which may be helpful in making sound-meaning-connections when exposed to previously unheard but well-formed utterances.

Method

The informants were randomly selected from the National Swedish address register (SPAR) on the basis of age and geographical criteria. A total of 55 infants, recorded during a period of 1 month, participated in this study. Six infants were excluded due to interrupted recordings (infant crying or not wanting to look at the film). The 49 infants left were randomly assigned to watch one (two or all three) of the films: 22 to Svensiska 1 (12ml, 10 fm, mean age 1;1), 20 to Svensiska 2 (12ml, 8fm, mean age 1;1), and 20 to Svensiska 3 (7ml, 13fm, mean age 0;11). The parents were not compensated for their participation.

'Visual Preference' paradigm – a version of Preferential Listening Procedure (Fernald, 1985) – was used to test infant perception. The infant was seated on the parent's lap. A television monitor showed a film and played the audio corresponding to the objects in the film. A video camera recorded the infant watching the film. The parent was wearing air-tight headphones (listening to music) through the whole procedure. The experimenter, located in next door, was able to via a monitor connected to the video camera look at the section taped. The informants looking behavior was measured with precision of 0.04s, half of them by an observer who performed 'blind' analysis, and the other half by an observer who knew the test condition assignment of the infants analyzed [1].

The films were divided in four sections [2]: **Part a** (30s) showed a split-screen with two objects: a boy and a girl doll, one of them to the left (L), the other to the right (R). The audio of this part was silent. **Part b** (60s) showed just one of the dolls. The audio was made of nonsense phrases, which repeated the target-word to be connected with the doll. **Part c** (60s) showed the other doll (alone) and played the nonsense phrases, which repeated the target-word to be connected with this doll. **Part d** (30s) showed, just like part a, a split-screen with the two objects, one of them to the L, the other to the R. The audio was made of nonsense questions like **Fur bu skrett X*? The *X* stands for a target-word referring to one of the dolls.

Target-word frequency rate and number of phrase final target-words in the high variability condition (Svensiska 1) and the low variability conditions (Svensiska 2 and 3 containing the same nonsense phrases) are seen in Table 1. In the high variability condition the target-word frequency rate was low (23 instances out of 212 words). About half of these target-words were in phrase final position (11 out of 23 target-words) and the other half in other than

phrase final position (12 out of 23 target-words). In the low variability condition the target-word frequency rate was high (34 instances out of 197 words). These target-words were always in phrase final position (34 out of 34 target-words). In sum, the probability for a word to be a target-word was high (0.17) in the low variability condition, and low (0.11) in the high variability condition. As well was the probability for a phrase to contain a target-word higher (1.0) in the low than in the high variability condition (0.71).

Table 1. Target-word frequency rate and number of phrase final target-words per test condition

Test Condition	Target-word Frequency Rate	Phrase-final Target-words
High Variability (Svensiska 1)	23 target-words/ tot 212 words	11 phrase final/ tot 23 target-words
Low Variability (Svensiska 2 and 3)	34 target-words/ tot 197 words	34 phrase final/ tot 34 target-words

Results

9 of 22 informants in the high variability condition had longer orientation times at the correct object (11 had longer orientation times at the incorrect object, and 2 exactly the same score at the correct and the incorrect object). In the low variability condition 28 of 40 informants had longer orientation times at the correct object. In addition, analysis on infants looking behavior, prior to (Pre exposure), and after exposure to the video-audio pairs (Post exposure) was performed (Figure 1). Pre exposure was measured in % as looking time at the *to-be-correct* object of total looking time during part a. Post exposure was measured in % as looking time at the *correct* object of total looking time during part d.

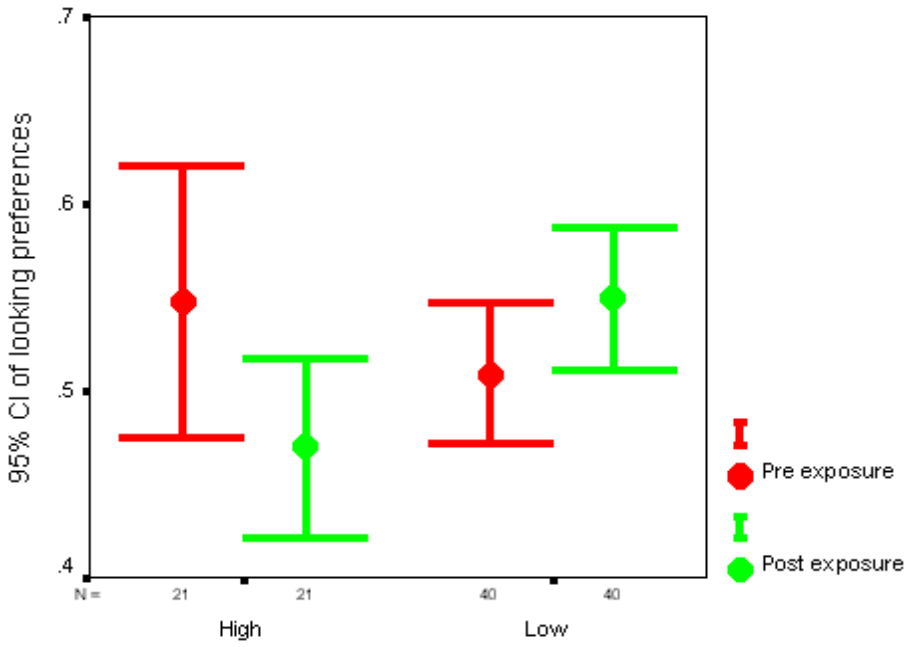

Figure 1. Y-axis: Change of looking preferences (from Pre to Post exposure), X-axis: Variability in speech sample. Since one of the infants had no exposure to part d, the number of informants in the high variability condition is (N21).

The informants in the high variability condition (N21) showed decreased (shorter Post exposure than Pre exposure) mean looking preference at the correct object, and the informants in the low variability condition (N40) showed increased (longer Post exposure than Pre exposure) mean looking preference at the correct object. An ANOVA revealed significant difference among the set of group means ($F= 7.749$; $df = 1/49$; $p<0.007$).

Summary and Conclusions

Recent work on infant speech perception has focused on identifying the speech cues which infants rely upon in order to identify where one word ends and the next begins. Research focused on statistics and algebraic rules as potential sources of information that listeners use in word segmentation were discussed. The general aim of the present study was to explore if babies manage to extract and associate novel target-words to novel objects. In addition it was suggested that statistical learning mechanisms might support even more complicated aspects of language learning such as acquisition of syntax.

Results showed that the majority of the informants in the low variability condition had longer orientation times at the correct object. In addition, in this condition infants looking preferences from the to-be-correct object to the correct object changed in 'right' direction. This supports the hypothesis that the nonsense phrases in the low variability condition (target-word present in each phrase, always in phrase final position) could 'explain' the objects just like semantically meaningful phrases do and facilitate in establishing association to the correct object.

The conclusion of *Marcus et al.* that infants learned rules governing sequences' grammar rested on the fact that the sequences contained novel syllables. However the 'rules' extracted might have created other statistical regularities - e.g. the AAB 'grammar' indicated that a syllable was followed by another instance of the same syllable and then a different syllable. To speculate about the nature of phenomena in the present study: The infants could take advantage of grammar – they were capable of recognizing stems of words as embedded words in derivated, inflected, and compounded words. The naturally produced nonsense phrases, which contained pauses and intonational patterns, supplemented the statistics inherent in the structure of phrases. In natural settings infants presumably benefit these other cues correlated with statistical information when learning different aspects of language.

Notes

1. The films varied in whether the girl or boy doll was shown to the L or R, or took turns in being presented as the first or second object. This was to control possible L/R preferences, and sequence influence of the objects.
2. As the measures of looking behavior by the two observers were in all essential aspects the same, the threat of experimenter expectancies was considered to be insignificant.

Acknowledgements

This study was carried out as Bachelor's thesis, supported by HSFR, and Telefon-aktiebolaget L M Ericsson. I thank my supervisor Francisco Lacerda for all kinds of support.

References

Altmann G.T.M. & Dienes Z. (1995): in Gao S.-J., Smith L.S., Hanckock P.J.B. (eds.) *Neural Computation and Psychology*, Springer-Verlag, New York.

Fernald A (1985): 'Four-month-old infants prefer to listen to motherese', *Infant Behavior and Development*, 8.

Gomez R.L. & Gerken L.-A. (1997): in Hughes E., Hughes M., Greenhill A. (eds.) *Boston University Conference on Language Development* 21, Cascadilla Press, MA.

Marcus G., Vijayan S., Bandi Rao S. & Vishton P. (1999): 'Rule Learning by Seven Month-Old Infants', *Science* 283.

Saffran J.R., Aslin R.N. & Newport E.L. (1996): 'Statistical Learning by 8-Month-Old Infants.', *Science* 274, 1926-1928.