THE UNINTELLIGIBILITY OF SPEECH TO CHILDREN: EFFECTS OF REFERENT Availability

E. G. Bard and A. H. Anderson

Human Communication Research Centre,
Department of Linguistics, Centre for Cognitive Science, University of Edinburgh
Department of Psychology, University of Glasgow

ABSTRACT
Speech addressed to children is said to be helpful in several ways, but redundant words in speech to adults tend to lose intelligibility [7-8]. Word tokens extracted from the spontaneous speech of the parents of 22 to 36-month-old children and presented in isolation to adult listeners show less intelligibility when the words are redundant in two senses: they occur in repetitions of an utterance (Experiment 1) or they refer to an entity which is physically present at the time of speaking (Experiment 2). These findings help to explain why word tokens randomly selected from speech to young children are less intelligible than those from speech to adults [2]. Because these words are difficult to recognize, they appear to induce child listeners to rely on the word's extra-linguistic context during the recognition process [1], such as adults are induced to rely on discourse context [3, 6].

1. INTRODUCTION
Children perform a remarkable bootstrapping operation when they simultaneously learn syntax and vocabulary by listening to running speech. Word tokens in spontaneous speech are so different from their citation forms that they have about a 50% chance of being recognized in isolation by adult listeners who share the speaker's vocabulary [10]. Given that the child's interpretation of linguistic context may be too incomplete to aid word recognition in all cases, categorization of non-canonical tokens as belonging to a particular word type or learning more about the structure of a language from strings of such tokens must be especially difficult.

The perceptual task might be simplified if parenthetically spoken words more clearly related to children than to adults, but on the contrary, words randomly selected from parents' speech to children (hereafter "A-C speech") aged 22 to 36 months proved significantly less intelligible out of context than words from the same parents' speech to an adult (hereafter "A-A speech") [2]. Alternatively, the well attested redundancy of speech to small children [9] may make their task easier. Words are more predictable from their sentence contexts in A-C Speech than in A-A speech [2]. Utterances to children are more often partly or completely repeated [9, 11]. A-C speech is also more supported by physical context, since it refers almost exclusively to objects and situations which are available to the child's senses at the time [9]. Perhaps some combination of the surrounding context, earlier occurrences of the same utterance and the physical presence of referent objects can be exploited by the child.

In A-A speech, however, redundant word tokens, both those more predictable from sentence context [7-8], and those referring repeatedly to the same entity [4-6], are shorter and less intelligible when isolated than their less redundant counterparts. If the effect applies for all kinds of redundancy, then words naming salient visible objects may also be less clear. In A-C speech, increased intelligibility from sentence context has been found to correlate with lowered word intelligibility [2]. This paper asks whether intelligibility also falls when A-C words refer to entities in the same way as objects.

2. EXPERIMENT 1: REPETITION
Experiment 1 tests the hypothesis that words spoken more than once in A-C utterances produced in close succession will be less intelligible than words in the first.

2.1. Method
Corpus. The materials were drawn from 12 45-minute studio-recorded sessions, in which each parent spoke to his or her child and to an experimenter. Both parents of one boy and one girl in each of three age groups (22-24 months, 28-30 months, 34-36 months) participated. After discussing with the parent the family's history and details of the child's contacts and play habits, the experimenter encouraged the parent to let the child play with a standard set of toys so that the child's speech in play might be recorded. The parent later labeled the child in conversation about one of his or her own toys which resembled one in the studio. Parent and child were recorded on separate channels of a Revox A77 stereo tape recorder, the parent via a lavalier microphone. Other details will be found in [2].

Tapes were fully transcribed in the standard orthography and all nouns spoken by the parents, except proper names, were classified according to the address and the location of the entity referred to. Present-noun names objects or persons in the studio which were being discussed or acted on by speaker and listener. Present-for-Speaker nouns referred to objects to which the speaker was thereby directing the listener's attention. Absent nouns referred to entities or events not present in the studio. Unclassifiable nouns referred to abstractions and to physical or topological entities in which the studio was contained. Table 1 summarizes the different distributions of A-C and A-A nouns among these categories.

Materials and Design. From the speech of each parent to his or her child, 4 pairs of word tokens were chosen. Each pair included two successive co-referential tokens of a single noun which occurred in self-repetition during two utterances in the same conversational turn, the second of which either exactly repeated or closely paraphrased the first without altering the noun phrase containing the selected word. Two pairs from each parent were Child-Present words, two Child-Absent.

Table 1. Distribution of Common Nouns from Speech of Twelve Parents by Addressee and Location of Referent (N = 4013; r^2 = 1371, df = 3, p < .0001)

<table>
<thead>
<tr>
<th>ADDRESSER</th>
<th>REFERENCE LOCATION</th>
<th>PRESENT</th>
<th>PRET FOR SPEAKER</th>
<th>UNCLASS.</th>
<th>ABSENT</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHILD</td>
<td>(%)</td>
<td>1587</td>
<td>(62)</td>
<td>495</td>
<td>406</td>
<td>2588</td>
</tr>
<tr>
<td>ADULT</td>
<td>(%)</td>
<td>70</td>
<td>(5)</td>
<td>583</td>
<td>785</td>
<td>1445</td>
</tr>
</tbody>
</table>

The selected items were excerpted from their taped contexts electronically and distributed among four groups to give balanced representation of speaker, token, and location. No group contained more than a number of one per parent. Each was presented in random order interspersed with materials from Experiment 2. Intensity levels were held constant as far as possible. Each word was preceded by a spoken number and repeated three times at approximately 5sec intervals.

2.2. Results
Figure 1 summarizes the results. The number of letter perfect or fully homophonic identifications of the stimulus showed the expected effect of Token: first tokens were more intelligible than second tokens (57.5% v 43.5%): F1 = 11.84, df = 1, 22, p < 0.05; F2 = 3.92, df = 1, 44, p < .05, Min F* = 2.94, df = 1, 64, .05 < p < .10. Thus, A-C speech shares with A-A speech a tendency to lose in clarity what it gains in repetitiveness [4-6].

3. EXPERIMENT 2: LOCATION
Table 1 illustrates a typical asymmetry between A-A speech, which refers largely to absent entities, and A-C speech, which deals with visible things. Even when first mentioned, however, Present nouns already "given" by extralinguistic context. The other location categories may include mentioned nouns which introduce new items. If linguistic and extra-linguistic contexts work similarly, then Present nouns should resemble co-referential or Given second mention in having relatively low intelligibility [4-6], whereas other categories
will include more intelligible words. The overall intelligibility difference between A-C and A-A speech might be partly due to the typical referent location for both, and should be lost if this factor is controlled.

3.1. Method

The corpus allowed balanced sampling from each parent only in Child-Present, Child-Unclassifiable, Child-Absent, Adult-Unclassifiable, and Adult-Absent categories. From each of these 4 tokens per parent were randomly selected. The 240 word tokens were prepared by the method described earlier and presented with the 96 tokens of Experiment 1 to the same 24 Subjects.

3.2. Results

Figure 2 shows the means for the 5 cells. Among the A-C words, the predicted effect of location was found: nouns with Absent referents were significantly more intelligible (65% correct recognitions) than those with Unclassifiable (45%) or Present referents (49%), while the latter did not differ significantly: one-way ANOVAs for Referent Location gave $F_2 = 29.14, df = 2, .05$; $t = 1.99, df = 2, 132, p < .05$; $Min. F$ n. s., Scheffé tests at $.05$.

For words to both Addressees, Unclassifiable nouns were less clear (49% correct) than Absent (62.5%), though the difference was significant only for words spoken to children: $t = 2.62, df = 1, 193, p < .05$; Scheffé tests by Subjects at $p < .05$. Since neither the Addressee effect nor the interaction was significant, there was no intelligibility difference due to Addresser alone.

4. GENERAL DISCUSSION

SOURCES OF REDUNDANCY

The redundant speech to children includes not just speech with reduced repeated words, but also, and more significantly, speech which is less intelligible and more prone to self-repetition. In general, children seem to be placing their young children at a disadvantage. To see how children might actually profit from these differences in the context of considering the uses to which adults put reduced repeated word tokens, Fowler and Housum [6] have shown that second tokens are better prompts to the recall of words associated in discourse with first tokens than are the first tokens themselves. They propose that the reduced second tokens signal reference to earlier material and so evoke the associated word. Alternatively, the process of recognizing the less intelligible second tokens may rely more heavily on linguistic context, merely evoking a representation of that context. To behave like adults, children would have to map less intelligible tokens onto known items while failing to do this for more intelligible words. Of the Addressee effect and the toys available, but in one condition and the other, the children were asked to recognize words excerpted from the present corpus [11]. The children were always familiar with the referent from a self-repetition available, but in one condition they could see the toys the puppet 'spoke', while in the other the toys were concealed in a box. Like the adult listeners in Experiment 2, the children found originally Absent words easier to recognize (59%) than originally Present words (45%) overall. In the case regardless of original addresser. Moreover, originally Present words were more readily identified when the toys were visible than when they were hidden (77%-59% correct, $N = 17, v = 36, N = 33, B = .279, t = 1.99, df = 48, p = .05$), whereas originally Absent words were less accurately identified when the toys were visible than when they were hidden (51%, $N = 30, 76$, $N = 14, \bar{B} = .362, t = -2.45$, $df = 44, p = .019$). Since children knew that all toys would be hidden or all would be visible in a given session, word pronunciation did not signal referent location. In the second condition children appeared to profit from the visible context to decode uninformative Present words, while that context proved a distraction when they attempted to decode the more intelligible Absent words. If these children performed in a typical way, then the unintelligibility of A-C speech encourages them to use supporting context in the process of recognizing what has been said to them. It is fortunate that this context is so often pertinent.

REFERENCES


This work was supported by an ESRC IRC grant to the University of Edinburgh and by ESRC Project Grant (SSRC HBI/61) to John Laver and the first author. The authors are grateful for the help provided by the participants in the study. Recent requests should be sent to the first author at the Human Communication Research Centre, University of Edinburgh, 2 Biscacle Place, Edinburgh EH8 9LL, U.K.

5. REFERENCES