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 supposed to be helpfully redundant 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 loss of 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 tokens are difficult to recognize, they appear to induce child listeners to rely on the word's extra-linguistic context during the recognition process [1], much 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 often 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, categorizing 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 parents habitually spoke more clearly 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 sentence, earlier occurrences of the same utterance and the physical presence of referent objects can be exploited by the child.

In A-A speech, however, more 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 predictability 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 just mentioned entities or denote physically present objects.

2. EXPERIMENT 1: REPETITION

Experiment 1 tests the hypothesis that words in the second of two nearly identical 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 a parent spoke to his or her child and to an experimenter. Both parents of one boy and

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

ADDRESSEE	REFERENT LOCATION				TOTAL
	PRESENT	PRES'T FOR SPEAKER	UNCLASS.	ABSENT	
CHILD	1587	80	495	406	2568
(%)	(62)	(3)	(19)	(16)	
ADULT	70	7	583	785	1445
(%)	(5)	(.5)	(40)	(54)	

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 help the child play with a standard set of toys so that the child's speech in play might be recorded. The parent later engaged 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 laveliere 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 addressee and the location of the entity referred to. *Present* nouns named 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 geographical 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, that is, in a pair of 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.*

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 one member of a word pair. 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.

Subjects and Procedure. Twenty-four native-speakers of English (6 per group) from the Edinburgh University community heard stimuli presented monaurally on a Revox A77. They were told that each stimulus was a word taken from conversational speech which they were to identify, by guessing if necessary.

2.2. Results

Figure 1 summarizes the results. The number of letter perfect or fully homophonous identifications of the stimulus showed the expected effect of Token: first tokens were more intelligible than second tokens (57.5% v 43.5%): $F_1 = 11.84$, $df = 1, 22$, $p < .005$; $F_2 = 3.92$, $df = 1, 44$, $p < .05$; $Min\ F' = 2.94$, $df = 1, 64$, $.05 < p < .10$. 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 words are already 'Given' by extralinguistic context. The other location categories may include mentions which introduce New items. If linguistic and extra-linguistic contexts work similarly, then Present nouns should resemble co-referential or Given second mentions 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 each, 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 (43%) or Present referents (49%), while the latter did not differ significantly: one-way ANOVAs for Referent Location gave $F_1 = 29.14$, $df = 2, 40$, $p < .005$; $F_2 = 3.29$, $df = 2, 132$, $p < .05$; *Min F'* n. s., Scheffé tests at $p < .05$.

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: *Min F'* = 4.08, $df = 1, 193$, $p < .05$; Scheffé tests by Subjects at $p < .05$. Since neither the Addressee effect nor the interaction was significant, there was no intelligibility difference due to Addressee alone.

4. GENERAL DISCUSSION

Sources of redundancy in speech to small children have a price. When an utterance is repeated, its words become less intelligible. When the objects spoken of are present to the senses, the referring nouns are also less intelligible. The effect cannot be attributed to occasional lapses in generally clear speech, for no A-C cell provides significantly clearer word tokens than the A-A cells (*e.g.*, Child-Absent at 65% vs. Adult-Absent at 60%). When redundancy rises, as in second tokens of repeated words (43.5%) or in Child-Present words (49%), intelligibility falls to below A-A levels. Even the Child-Unclassifiable words patterned like the unintelligible group (43%) while the Adult-Unclassifiable patterned like the clearer Absent cells (55%). Whatever the internal breakdown of the Unclassifiable cells may be, they do nothing to maintain high intelligibility for speech to children.

Of the A-C figures, the lower ones must be taken as typical. Although the present corpus is not a random sample of conversation types, it resembles those in other studies [11]: Child-Present words, which should be relatively unintelligible, predominate, while the clearer Child-Absent words were relatively rare, even when parents were instructed parents to produce them. Although self-repetition is not so common in A-C speech as Present reference, it is certainly more typical here than in adult conversation [11]. Consequently, the differences in intelligibility of large random samples of A-A and A-C speech [2] may have something to do with the tendency of

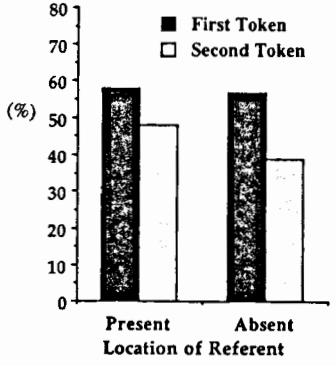

Figure 1. Intelligibility of Words in Repeated Utterances to Children (N = 96)

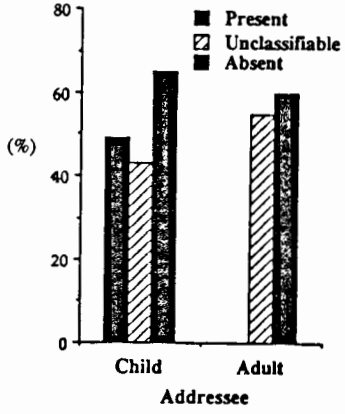

Figure 2. Intelligibility of Words by Addressee and Referent Location

speech to children to provide the expensive forms of redundancy which have been explored here. Certainly the difference is lost when referent location is held constant.

By succumbing to processes which reduce clarity when contextual support is high, parents seem to be placing their young children at a disadvantage. To see how children might actually profit from these difficulties, it is worth 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 lingistic context, thereby reactivating a representation of that context [3]. To behave like adults, children would have to map less intelligible tokens onto known items while failing to do this for more intelligible words.

Exactly this result was found when three-year-olds were asked to fetch the toys a puppet requested via tape recordings of the words excerpted from the present corpus [1]. The children were always familiar with the nouns used and the toys available, but in one condition they could see the toys as 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 this case regardless of original addressee. Moreover, originally Present words were more readily identified when the toys were visible than when they were hidden (63% correct, $N = 17$, v 36%, $N = 33$; $\beta = .279$, $t = 1.99$, $df = 48$, $p = .05$), whereas originally Absent words were less accurately identified when the toys were visible than when they were hidden (51%, $N = 30$, v 76%, $N = 14$; $\beta = -.362$, $t = -2.45$, $df = 44$, $p = .019$). 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'. Instead children appeared to profit from visible context to decode unintelligible 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.

This work was supported by an ESRC IRC grant to the University of Edinburgh and by ESRC Project Grant (SSRC HR6130) to John Laver and the first author. The authors are grateful for the help provided by the participants in the study. Reprint requests should be sent to the first author at the Human Communication Research Centre, University of Edinburgh, 2 Buccleuch Place, Edinburgh EH8 9LL, U.K.

5. REFERENCES

[1] BARD, E. (1982), *The Intelligibility of Speech Addressed to Children*, Unpublished PhD Dissertation, University of Edinburgh.
[2] BARD, E., & ANDERSON, A. (1983), "The unintelligibility of speech to children", *JChLang, 10*, 265-92.
[3] BARD, E., BREW, C., & COOPER, L. (1991), "Psycholinguistic Studies on the Incremental Recognition of Speech: An Introduction to the Messy and the Sticky". Deliverable R1.3.A, DYANA, ESPRIT Basic Research Action BR3175
[4] BARD, E., LOWE, A., & ALTMANN, G. (1989), "The effects of repetition on words in recorded dictations", *Proc. EUROSPEECH '89, 2*, 573-6.
[5] FOWLER, C. (1988), "Differential shortening of repeated content words produced in various communicative contexts", *LangandSpeech, 28*, 47-56.
[6] FOWLER, C. & HOUSUM, J. (1987), "Talkers' signaling of 'new' and 'old' words in speech and listeners' perception and use of the distinction", *JML, 26*, 489-504.
[7] HUNNICUTT, S. (1985), "Intelligibility vs redundancy - conditions of dependency", *LangandSpeech, 28*, 47-56.
[8] LIEBERMAN, P. (1963), "Some effects of the semantic and grammatical context on the production and perception of speech", *LangandSpeech, 6*, 172-5.
[9] NEWPORT, E., GLEITMAN, H., & GLEITMAN, L. (1977), "Mother, I'd rather do it myself: some effects and non-effects of maternal speech style", In C. Snow and C. Ferguson (eds.), *Talking to Children*, Cambridge: CUP.
[10] POLLACK, I. & PICKETT, J. (1963), "The intelligibility of excerpts from conversation", *LangandSpeech, 6*, 165-171.
[11] SNOW, C. (1972), "Mothers' speech to children learning language", *ChDev, 43*, 549-65.