

Perceptual restoration in toddlers

ROCHELLE S. NEWMAN

University of Maryland, College Park, Maryland

Perception by adults is a constant interaction between the top-down effects of prior knowledge and the effects of bottom-up perceptual information. One obvious example of this interaction is the perceptual restoration effect, in which adult listeners have the illusion that a word is complete when a portion of it has been replaced by a masking noise. In four experiments, we demonstrate that toddlers fail to show the illusion of perceptual restoration, even in constrained situations with words they know quite well. Not only do toddlers have less prior knowledge than do adults, but they also appear to place less reliance on the knowledge that they do have, at least in the paradigm tested here. Instead, toddlers appear to be more tied to the perceptual information they receive than are adults.

We often listen to speech in situations that are not perfectly quiet. Other talkers, machine noises, even biological noises such as coughs and stomach gurgles can mask a talker's speech temporarily. Yet even when an outside noise masks part of the incoming speech signal, we generally have little difficulty interpreting the intended message. Our prior knowledge of the language allows us to determine what the speaker intended to produce, even when part of the perceptual information is blocked.

Listening to speech in noise can also be a problem for young children (Newman, 2005). They may have to listen to their caretakers' speech in the context of noise from their siblings playing, from the television down the hall, and from the cars on the street outside. Yet toddlers do not have the same background knowledge that adults have. They may not be able to rely on knowledge of the words

in the language or of semantic context to help them figure out intended meanings. Thus, toddlers may be faced with an even greater difficulty in listening to language than are adults, at a time when they are simultaneously trying to learn the language.

When do toddlers begin using their developing pool of language knowledge to aid in their interpretations? One possibility is that the use of this information is mandatory in language processing. If so, children will be affected by such knowledge to the extent that they have it and, thus, will gradually begin showing stronger top-down language effects as they learn more about the language. According to this idea, the reliance on prior knowledge is constant throughout development; the only change with time is the extent to which this prior knowledge exists.

Another possibility is that the use of such knowledge itself develops over time. Not only do toddlers have less prior knowledge, according to this view, but also they rely less on the partial knowledge they do have. Perhaps a critical amount of such knowledge is required before the language-processing system will depend on it. This would suggest that there is a reorganization at some point (or points) during development in which greater emphasis is placed on top-down knowledge.

The present set of experiments was performed to examine whether toddlers make use of one such form of prior knowledge typically used by adult listeners, shown in the phoneme restoration effect (Samuel, 1996b; Warren, 1970; Warren & Obusek, 1971; Warren & Warren, 1970). Warren presented listeners with the sentence, "The state governors met with their respective legislatures convening in the capital city," but he replaced the first "s" in legislatures with an extraneous noise (either a cough or a tone). He found that listeners not only reported hearing the word "legislatures," but also could not accurately report the location of the noise in the sentence. That is, listeners actually heard the interrupted word as continuing "behind" the noise. This is in stark contrast to the results when a phoneme is replaced with silence: Although listeners can tell what the word should have been in this case, they do

Some of the present results were reported at the 139th meeting of the Acoustical Society of America in Atlanta, June 2000. This work was supported by Research Grant HD37822-01 from NICHD and Research Grant BCS 99-07849 from NSF to the University of Iowa. The author thanks Lloyd Frei for the design of the video switching mechanism, Keith Miller and Jim Garmon for the implementation of the test booth/subject-running system, Ian Palmer and Don Mathis for programming the testing system, and Lisa Oakes for helpful discussions. Jennifer Alexander, Christine Beagle, Marianna Christodoulou, Emilie Clingerman, Sheryl Clouse, Nicole Craver, Hayley Derris, Rachael Dolezal, Annie Ferruggiaro, Meredith Good, Sarah Haszko, Ryan Hurm, Isma Hussain, Lacey Kahner, Hannah Kim, Lisa Loder, Heather McIntosh, Robin Nicoletti, Jessica Pecora, Rebecca Ribar, Amy Robbins, Antonia Rodriguez, Kate Shapiro, Lauren Simpson, Emily Singer, Sarah Stilwill, Ben Schnoor, Phillippe Taborga, Cheryl Tarbous, Susan Timm, Andrea Tuttle, Suji Uhm, Stephanie Weinberg, Tammy Weppelman, Brooke Werner, SangEun Woo, Tyler Wunnenberg, Donnia Zack-Williams, and Jenni Zabler assisted in subject running and/or coding. Thanks also to Peter Jusczyk for the alternative interpretation of Experiment 1 involving cross-trial comparisons, which led to Experiment 2, and to Letitia Naigles, Robert Remez, Arty Samuel, and two anonymous reviewers for previous comments on the manuscript. Finally, my thanks to Jodie Plumert and Tim Barrett for allowing me to tape their dog, and to Sheryl and Kelly Clouse for videotaping their cat. Correspondence should be addressed to R. S. Newman, Department of Hearing and Speech Sciences, University of Maryland, College Park, MD 20742 (e-mail: rnewman@hesp.umd.edu).

not have the illusion of the word's being complete (Warren & Obusek, 1971). Whereas noise provides an explanation for why the expected sound was not actually heard (it was masked), silence does not, and thus illusory restoration does not occur.

Subsequent work has compared adult listeners' perception of noise-replaced stimuli (where the phoneme in question is excised before noise is added) with the perception of noise-added stimuli (where the noise is superimposed on the existing phoneme). With this procedure, Samuel (1981a) was able to distinguish between subjects' biases and their discrimination abilities (in a signal detection analysis). He demonstrated that the restoration effect involves more than simply a bias in responding; rather, prior knowledge about words changes the listeners' perception (listeners actually hear words as being intact more often than nonwords; see also Samuel, 1996a, 1997).

There are clear limits on the generality of perceptual restoration. It occurs only when the noise is sufficiently similar to the acoustic signal of the missing speech sound as to provide a reasonable explanation for why the speech sound was not heard. That is, restoration can occur only when an extraneous sound is capable of masking a signal. Thus, in Warren's (1970) original example, the phoneme /s/ contains energy at a wide range of frequencies; a cough contains energy at many of these same frequencies and thus, logically, can be a masker. Restoration would not have occurred had that same phoneme been replaced with a single tone. Phoneme restoration is thus contingent both on the class of the phoneme being masked and on the acoustic properties of the masker (Samuel, 1981a); only when these are somewhat similar does the effect take place.

Despite the fragile nature of the effect, the fact that it occurs at all suggests that top-down cognitive knowledge can influence speech perception, at least in some situations. Moreover, this restoration effect is not limited to speech or to single phonemes. Restoration has been reported for other domains, such as music (DeWitt & Samuel, 1990; Sasaki, 1980), other modalities, such as American Sign Language (Schultz-Westre, 1985), and other species, such as starlings (Braaten & Leary, 1999). It appears to be one example of a more general class of restorative events known as auditory induction (Bashford & Warren, 1987).

This effect has also been examined in children 5 to 6 years of age (Koroleva, Kashina, Sakhnovskaya, & Shurgaya, 1991; Newman, 2004; Walley, 1987). Children this age appear to be more affected by signal disruptions than are adults. Regardless of the type of disruption (addition of noise, substitution of noise, or substitution of silence), children appear less able to rely on other forms of knowledge to fill in the missing information (Walley, 1987). Despite this, children are better able to identify words when phonemes are replaced with noise than when they are replaced by silence (Newman, 2004). Although children are more sensitive to disruptions in general, their perceptual restoration appears to have already developed by the time they reach 5 years of age.

No studies to date have examined restoration in younger children, whose lexicons are still in a state of rapid development. Obviously, perceptual restoration can occur only in words known to the listener (or primed by the task; see Samuel, 1981a). But children begin to recognize their first words as young as 6 months of age (Tincoff & Jusczyk, 1999) and have a fairly substantial vocabulary by the time of their second birthday (Dale & Fenson, 1996; Fenson et al., 1994). This lexical knowledge could potentially be used during the course of language processing. The present study was designed to investigate this issue. More specifically, we investigated whether toddlers would be more likely to identify a known word when a portion of the word was replaced with noise, rather than with silence.

In these experiments, we presented children with words that either were fully intact or had a phoneme replaced either by noise or by silence. Children's identification of these words was examined by measuring the amount of time they spent looking at an appropriate versus an inappropriate referent. The critical comparison was between the situation in which a sound was missing but there was no potential masker present and the condition in which a noise served as a potential masker for the missing sound. If the children performed better in the noise condition than in the silence condition, this would be an indication that noise enhanced the intelligibility of the signal. If there was no difference, it would suggest that the children were not using their prior knowledge of the lexicon and of potential maskers to fill in missing information.

We did not necessarily expect performance in the noise condition to reach that in the clear condition; restored phonemes may not reach the same level of identification as do intact words. However, throughout this article, we will compare performance in the altered conditions with that in the clear condition as a check for ceiling effects. If the noise and silence conditions do not differ from the clear condition, it might be an indication that the task was too easy; we would not expect to be able to find differences between the noise and silence conditions if performance in both was at ceiling levels (equivalent to that in the clear condition).

EXPERIMENT 1

The present experiment was designed to examine whether toddlers would show perceptual restoration when presented with a known word. We selected two words that children typically learn quite early: "kitty" and "doggie." Both of the referents for these words are easily picturable and are likely to be approximately equally interesting to toddlers as a group. Moreover, both words contain a stop consonant as a medial phoneme, which Samuel (1981a, 1981b, 1996a) has shown to be a class of phonemes particularly prone to restoration.

Toddlers were tested using the preferential-looking paradigm (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987). They were presented with two video screens, one showing a cat, the other showing a dog. They heard sentences containing the two target words in each of three conditions:

clear (original unaltered productions), noise (the medial stop consonant of the target word was removed and replaced with noise), and silence (the medial stop consonant was removed and replaced with silence).

The clear condition served as a test of the children's word knowledge: If the toddlers knew the meanings of the two words, we would expect them to look longer at the cat picture when hearing the label "Kitty!" and to look longer at the dog picture when hearing the label "Doggie!" If the toddlers did not look appropriately in the clear condition, no further examination of the results would be warranted.

Assuming that the toddlers did look appropriately in the clear condition, we could then compare their results in the noise and the silence conditions. If the toddlers were able to use partial information to determine a word's meaning, they should look longer at the appropriate animal than at the opposite animal in both the noise and the silence conditions. With only two choices as to what to look at (a cat and a dog), even the partial information of "ki__y" and "do__y" should lead them to look longer at the appropriate video (Fernald, Swingley, & Pinto, 2001). Perceptual restoration would be demonstrated by a larger amount of appropriate looking when a phoneme was replaced with noise (which could theoretically be masking a phoneme that was actually present) than when a phoneme was replaced with silence.

It is worth noting that this task did not test specifically for the illusion of restoration (i.e., of the word's being heard as complete in the noise condition, but not in the silence condition). The presence of this illusion could have been tested only with explicit questioning, which was not appropriate for children of this age. Instead, we looked for an increase in the intelligibility of the word when the gaps in the signal were filled with a potential masker than when left unfilled. This method of testing for perceptual restoration had previously been used by a number of investigators (e.g., Bashford, Riener, & Warren, 1992; Bashford, Warren, & Brown, 1996).

Samuel (1981a) has suggested that comparing stimuli where noise replaces a phoneme with stimuli where noise is added to a phoneme is a more sensitive test of restoration than is comparison of noise-replaced and silence-replaced conditions. However, the difference between noise-replaced and noise-added stimuli is quite subtle, and finding such a difference is dependent on presenting a large number of trials. This is not feasible with toddlers. Moreover, results with young listeners tend to be far more variable than those with adult listeners. Between the more variable task performance and the much smaller number of test trials possible, it seems unlikely that differences between noise-added and noise-replaced stimuli could be detected with young children. For that reason, we chose to compare only noise-replaced versus silence-replaced conditions, while admitting that this may not be as sensitive a task as that performed with adult listeners in Samuel's studies.

Method

Subjects. Twenty 24-month-old children (12 of them male and 8 female; range, 100–109 weeks; mean, 104 weeks) participated

in this experiment.¹ All were developing normally, and none had evidence of an ear infection at the time of test. Data from an additional 8 toddlers were discarded for the following reasons: equipment failure/experimenter error (4), fussiness (2), ear infection (1), and outside distraction (1).

Stimuli. The visual stimuli were videotapes of a typical dog (a yellow Labrador) and a typical cat (a black shorthair). Both were relatively calm animals and remained centered on the videotape for the entire 10-min recording, with only small amounts of movement (such as tail wagging and self-cleaning).² Each animal was taped on a solid background of an easily discriminable color (a white bedspread for the black cat and a black sheet for the yellow dog).

The auditory stimuli were recorded by a native speaker of midwestern English, R.S.N. All items were recorded in a noise-reducing chamber at a 44.1-kHz sampling rate, 16-bit resolution, and were stored on computer disk. The speaker first recorded several tokens of the target items as clear exemplars. These items were used to ensure that the toddlers already knew the appropriate referents for these target words. Rather than using these clear tokens as the base for the restoration stimuli, however, we decided to use a more conservative approach, in which transitional cues for the missing phonemes were removed. This ensured that any coarticulatory cues to the missing phoneme did not remain in the signal once the phoneme was removed. New recordings were made in which the target words were mispronounced prior to recording. (Thus, rather than recording "kitty" and "doggie," recordings for the noise and silence trials consisted of such items as "kippy" and "dobby"; see Warren & Sherman, 1974, for an explanation of this method of production.) Cutpoints were selected on the basis of visual inspection and a spectral analysis of the recordings and were chosen so as to minimize any transitional cues to the medial stop consonant. (However, it is important to note that any cues remaining would have indicated an incorrect phoneme and, thus, would have been unlikely to lead to restoration.) The medial stop consonant was removed and replaced with either silence or a noise.

Pilot testing suggested that sinusoidal tones served as the best noise for the voiced stop consonant in "doggie" but that white noise worked best for the voiceless stop consonant in "kitty."³ Since we wanted to use the same noise for both target words but wanted the two items to be similar in their acoustic properties, we decided to blend these two noises together. The final noise consisted of white noise mixed with sinusoidal tones of 100, 200, 300, and 400 Hz. This noise was used for all of the noise condition stimuli.

We created a number of edited tokens of "kitty" and "doggie" and selected among them on the basis of pilot testing with adult subjects. Items were chosen that had the highest identification as "kitty" and "doggie" in the noise condition and the lowest such identification in the silence condition. These eight tokens of each word were reliably restored by adult listeners in our pilot testing, although they did not result in unanimous judgments. (This issue will be discussed further in the Discussion section and will be addressed in Experiment 3.) The eight tokens were appended to one another in a sequence, and this sequence was then presented on the test trials. The same tokens, in the same order, were then used for both the noise and the silence conditions.

Thus, on each test trial, the toddlers heard a woman's voice repeating the target word eight times ("Kitty! Kitty! Kitty!" etc.). On clear trials, all eight tokens were pronounced naturally and were unedited. On noise and silence trials, all eight tokens were ones that had been originally mispronounced and then edited to remove the medial consonant. It is important to note that the *identical* stimuli were used for the noise and the silence conditions. Transition information was identical in both cases, as were stimulus cutpoints. The only difference between these items was the presence of noise versus silence during the gap left by removing the medial stop consonant. Thus, these two conditions could be directly compared to determine the role of the noise masker on perceptual restoration.

Procedure. The toddlers were seated on their parents' laps in a three-sided booth enclosure. The front panel of the booth contained

two 20-in. video screens, separated by 18 in., and a center white light. This light flashed at the beginning of each trial to direct the toddlers' attention to the center of the booth (equidistant from the two videos). Trials were initiated only when the child was attending to the light; the experimenter then pressed a button on a computer-controlled response box, signaling the computer to turn off the light and begin the next trial. The auditory stimuli were played over matched NHT speakers, one located on each side of the enclosure.

The videotapes of the dog and the cat played continuously through the session, but the signal was routed to the monitor only during appropriate trials. Each monitor showed a basic black background when it was not presenting the video picture. The computer controlled the switch from black screen to video, thus ensuring that all the trials were of identical duration and that the video and audio signals were in synch.

During the first four trials, only one video was shown on each trial. These trials lasted 5 sec and were designed to ensure that the child learned which video monitor showed which picture. Which video appeared on which screen and which video was played first were counterbalanced across subjects.

The next four trials served as a baseline measure of the children's looking preference. Each trial lasted 10.7 sec (the same duration as the test trials). Both videos were played during these trials, but the toddlers did not hear any auditory stimuli. Thus, the toddlers' looking time during these trials serves as a measure of the toddlers' preference for one visual stimulus over another.

The following 12 trials served as test trials. The audio signal began 1 sec prior to the video, which lasted 10.7 sec on each trial. The toddlers heard each of the six trial types ("kitty" vs. "doggie," and clear vs. noise vs. silence) two times each. These were presented in two blocks, so that each of the six trial types occurred once in a randomized order in the first block and then each occurred in a new randomized order in a second block. Each child received a different random order of trials.

During the test session, the parent wore a visor hung with black cloth, preventing him or her from seeing the two video screens. Thus, the parent was not aware of which television was showing which image. The experimenters also watched the parents to ensure that they did not lift their visor during the session.

Sessions were videotaped, and two experimenters (blind to condition) individually coded each child's looking behaviors using Habit (Cohen, Atkinson, & Chaput, 2000). The averages of these two codings were used as the final data. Two measures were taken from each trial: the child's total duration of looking in each direction and the single longest look in each direction. Correlations between the two coders were .95 for overall durations and .94 for single longest looks. Average differences were less than 0.5 sec in both cases. Despite this high reliability, there were occasional trials on which the coders disagreed. On any trial in which the two coders disagreed by 2 sec or more, a third experimenter coded the same trial, and the two closest results were selected as the final scores for averaging.

Analysis. Several comparisons were calculated from the coding measures, and each was performed separately for the clear, noise, and silence conditions. The first comparison was on average duration of looking to the matching versus the mismatching video. (Use of this particular measure has been supported by Golinkoff et al., 1987.) The total number of seconds spent watching each of the two videos was averaged across test trials. This resulted in a single value (in seconds) for the amount of time the child spent watching the matching video (i.e., the cat on "kitty" trials and the dog on "doggie" trials) and a second value (also in seconds) for the amount of time spent watching the mismatching video (the cat on "doggie" trials and the dog on "kitty" trials). These values were compared using a paired *t* test.

The second measure calculated the percentage of time the toddler spent looking at the appropriate video minus the percentage of time spent watching that video in the baseline condition (i.e., what would be expected by chance). This value was then compared with zero in

a one-group *t* test. This measure was taken separately for each of the two videos. Thus, we had one test for the percentage of time the child spent watching the cat video in the "kitty" test condition versus the baseline condition and a second test for the percentage of time the child spent watching the dog video in the "doggie" trials versus the baseline trials. This allowed us to determine whether the overall preference for matching versus mismatching videos was consistent for both videos or was limited to only one of the two videos.

The third measure involved the single longest look on each trial, rather than the overall looking. It involved calculating the longest look on test trials minus the longest look on the baseline trials. This, too, was calculated separately for the two videos, and the resultant values were compared with zero in one-group *t* tests. Schafer and Plunkett (1998) have argued that the durations of the longest looks are a more sensitive measure of toddlers' preferences than are overall looking measures.

Since we predicted that toddlers would watch the matching videos longer than the mismatching videos, we had a directional prediction for all the measures. We therefore used one-tailed *t* tests in all cases.

Results and Discussion

Unless the toddlers looked longer at the appropriate video in the clear condition, no other analyses would have been warranted. We therefore began by analyzing these clear condition results. Toddlers watched the matching video 3.28 sec longer than the mismatching video in this condition, a significant difference [$t(19) = 6.95, p < .0001$]. This effect was not limited to only one of the two videos. Toddlers watched the cat video longer on "kitty" trials than on baseline trials [$t(19) = 2.95, p < .005$] and watched the dog video longer on the "doggie" trials than on baseline trials [$t(19) = 4.50, p < .0005$]. The longest look measures showed similar results: Single longest looks were longer on appropriate test trials than they were on baseline trials for both the cat [$t(19) = 4.49, p < .0005$] and the dog [$t(19) = 4.39, p < .0005$]. These results showed that the toddlers understood the meanings of the two words "kitty" and "doggie" and would watch the appropriate video when they recognized one of these words.⁴

Given these results from the clear condition, we were now in a position to compare the results for the other two conditions. The overall matching results for the three conditions can be seen in Figure 1; data from all the measures can be seen in Table 1.

For the silence condition, the toddlers watched the matching video longer than the mismatching video [$t(19) = 2.21, p < .05$], although the amount of this difference (1.15 sec) was much smaller than that in the clear condition. This effect was not significant in the percentage of total looking for either video individually [$t(19) = 0.58, p > .10$ for the cat; $t(19) = 1.44, p < .10$ for the dog]. However, the effect was present in the longest looks for both videos [$t(19) = 2.46, p < .05$ for the cat; $t(19) = 2.09, p < .05$ for the dog]. These results suggest that the toddlers could use the partial information in these recordings to help direct them toward the appropriate video. However, despite the fact that the toddlers had only two choices (i.e., when they heard the word "ki_y!") their only options were a cat or a dog), the toddlers did not perform as well in this condition as they did in the clear condition.

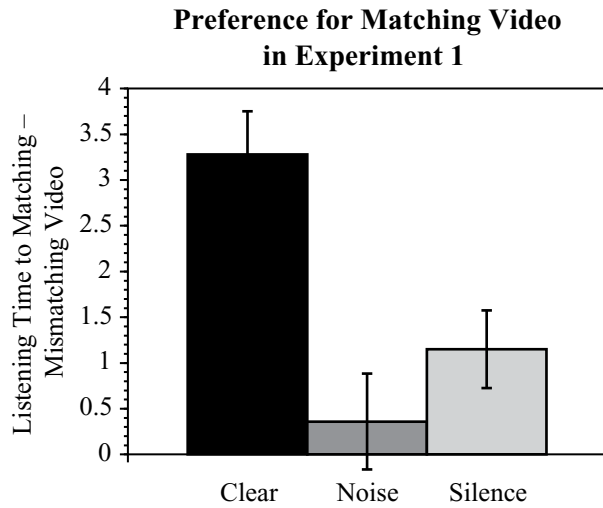


Figure 1. Differences between the time the subjects spent watching the matching video versus the mismatching video in Experiment 1.

Although the toddlers appeared capable of using partial information to direct their attention, they did not respond as strongly to this information as they did to speech presented in full detail.

We then examined the noise condition. The toddlers in this condition did not watch the matching video longer than the mismatching video [$t(19) = 0.85, p > .10$]. Nor did they demonstrate significantly longer looking to either video on test trials than in the baseline condition, although the effect for the dog video was marginal [$t(19) = -0.38, p > .10$ for the cat; $t(19) = 1.42, p < .10$ for the dog]. Similar results were shown with the longest look measure [$t(19) = 1.29, p > .10$ for the cat; $t(19) = 1.40, p < .10$ for the dog].

If the presence of the noise had resulted in increased intelligibility (or if they had experienced the illusion of perceptual restoration in the noise condition), the toddlers

presumably would have attended to the appropriate video longer in that condition than in the silence condition. This clearly did not occur. Indeed, the children appeared to show a greater degree of preference for the appropriate item in the silence condition than in the noise condition: Comparing across conditions, the preference for the matching over the mismatching video was significantly larger in the silence condition than in the noise condition [$t(19) = 2.56, p < .02$], and the toddlers' longest looks to both videos were larger in the silence condition than in the noise condition [$F(1,19) = 5.22, p < .05$]; there was no significant difference in percentage of looking [$F(1,19) = 1.63, p > .20$]. This opposite-direction effect suggests that the toddlers did not restore the missing phoneme.

Across the three conditions, 19 of the 20 toddlers looked longer to the appropriate video during the clear conditions. Only 12 looked longer at the appropriate video during the silence condition, and 10 did so in the noise condition. Although the children did appear to know these two target words, they did not use that prior knowledge differently in the noise condition than in the silence condition, nor did the presence of noise make the words themselves more intelligible.

There are several possible reasons for such a finding. One possibility, described earlier, is that the use of prior lexical knowledge develops over time. Not only do toddlers have less prior lexical knowledge, but also they rely less on the partial knowledge that they do have. Perhaps a critical amount of such knowledge is required before the language-processing system will depend on it. A related alternative is that children may remain more open to new words and are, thus, more likely to treat the altered sequences as being new words, rather than as being mispronunciations of already known words.

Another possibility is that lexical entries in general are less well known to children than they are to adults. Even though the children knew these words (as was shown by appropriate looking in the clear condition), their lexical

Table 1
Overall Looking Results in Experiment 1

	Condition					
	Clear		Silence		Noise	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Overall duration of looking (sec)						
To matching video	5.62	0.33	4.69	0.40	4.16	0.33
To mismatching video	2.34	0.19	3.54	0.25	3.80	0.25
Percentage of looking						
To cat, "kitty" trials	74.2	4.1	61.1	5.1	56.2	4.4
To dog, "doggie" trials	64.8	4.4	50.4	5.5	49.2	5.6
Longest looks (sec)						
To cat, "kitty" trials	5.06	0.49	3.92	0.47	3.32	0.38
To dog, "doggie" trials	4.17	0.43	3.33	0.51	2.95	0.50

Note—In all the statistics on percentages of looking and longest looks, the difference scores between the numbers above and the values obtained on baseline trials were examined. That is, what was crucial was not the number of seconds the cat was watched on "kitty" trials but whether there was an increase above that found on baseline trials. On baseline trials, overall percentages of looking were 58.3% to the cat ($SE = 3.6\%$) and 41.8% to the dog ($SE = 3.6\%$). Longest looks were 2.77 sec ($SE = 0.23$) to the cat and 2.22 sec ($SE = 0.23$) to the dog.

representations for the words may not have been strong enough to support restoration. Indeed, other research has suggested that there is an extended time period during which lexical representations develop (McGregor, Friedman, Reilly, & Newman, 2002). Although toddlers can recognize the words “kitty” and “doggie,” they may not have fully developed semantic representations for them.

Regardless of the reason, the toddlers did not appear to show the same pattern of restoration as that previously found with adult listeners. However, there are several possible problems that may have affected these results. The first potential difficulty is that the toddlers might have made cross-trial comparisons. Having heard some items with clear examples of the words “kitty” and “doggie,” the noise and silence conditions might simply have sounded relatively poor in comparison. This might have been especially the case for the noise items, since the loud noise might have sounded somewhat startling, in comparison with the normal speech conditions. It does not appear that the latter was the case, since the toddlers did not spend less time watching the videos overall in the noise conditions ($F < 1$).⁵ However, one way to ensure that these types of cross-trial comparisons do not occur would be to repeat the same experiment in a between-subjects manner: If toddlers heard only the noise or only the silence condition, there could be no comparison with the same words presented in the clear. This was the goal of Experiment 2.

The second possible problem is with the words themselves. Restoration effects are sensitive to phoneme class. Although Samuel (1981a, 1981b) demonstrated better restoration for stop consonants than for fricatives in most conditions, these results were obtained with adult listeners and may not carry over to children. Moreover, Trout and Poser (1990) suggested that the advantage for stop consonants might be due primarily to nonalveolar voiceless stop consonants. If this is the case, the voiced /g/ of “doggie” and the alveolar flap of “kitty” may not be ideal choices for restoration. Furthermore, stop consonants are partially cued by silence, especially word internally, so the silence condition may actually induce a stop consonant percept, making it sound more equivalent to (or even better than) the noise. (Neither Samuel, 1981b, nor Trout & Poser, 1990, used a pure silence condition in their experiments in which they showed strong restoration for stop consonants.) In addition, our adult listeners did not have unanimous judgments of these stimuli; some tokens resulted in the illusion of a complete word, but others did not. Although we are not focusing on the illusion of completeness per se, a much stronger case could still be made if we tested toddlers with items that adults restored more consistently. This issue was addressed in Experiment 3.

A third concern is the fact that the children heard only isolated words. Perhaps they did not take this series of words as an instruction to look at that particular item. Or perhaps the effect of restoration would be stronger in a sentential context. We decided to switch to using a sentential context for the remaining experiments, having the target words appear in phrases such as “Look at the ____!” and “Find the ____!”

EXPERIMENT 2

The toddlers in Experiment 1 failed to demonstrate evidence of perceptual restoration. Although they did show some evidence of the ability to use partial information to determine a speaker’s intended meaning, they did not show better performance in the noise condition than in the silence condition.

Yet the within-subjects nature of this experiment may have caused the toddlers to compare items across trials. Since the toddlers heard clear examples of the target words on some trials, they may have treated the experimental versions of these same words as poorer examples. Other research has demonstrated that even young infants are sensitive to very detailed information about target words (Fennell & Werker, 2003; Jusczyk & Aslin, 1995; Swingley & Aslin, 2002). Given this sensitivity to detailed information, the toddlers may have treated the noise and silence stimuli in Experiment 1 as having a different meaning than the clear versions of the same words.

The present experiment differed from Experiment 1 in two ways: It involved a between-subjects comparison, rather than a within-subjects comparison, and it presented the target words in a sentential context, rather than in isolation. The toddlers in this experiment were assigned randomly to hear the clear versions of both words, the noise-replaced versions of both words, or the silence-replaced versions of both words. Thus, there could be no comparison between poorer and better exemplars across trials. If the results in this experiment mirrored those in Experiment 1, it would suggest that cross-trial comparisons were not the cause for the toddlers’ lack of perceptual restoration.

Method

Subjects. Thirty-six 24-month-old toddlers (20 of them male and 16 female; average age, 105.2 weeks; range, from 101 weeks to 108 weeks, 1 day) participated in this experiment. Twelve toddlers were assigned randomly to each of three conditions: clear, noise, and silence. The three groups did not differ in average age (clear, 105 weeks; noise, 105.7 weeks; silence, 104.9 weeks; $F < 1$). All the toddlers were developing normally and were from homes where English was the primary language. Data from an additional 7 toddlers were discarded for equipment failure ($n = 2$), experimenter error ($n = 1$), fussiness ($n = 2$), or outside distraction ($n = 1$) or for suddenly moving outside of the range of the camera ($n = 1$).

Stimuli. The visual stimuli were identical to those used in Experiment 1. For the auditory stimuli, the target words themselves were acoustically identical to those in Experiment 1, but they were placed in a new context. The speaker from Experiment 1 recorded sentences containing the target words (“Look at the kitty! Can you find the kitty?” etc.). These sentential contexts were chosen so as to be uninformative as to the identity of the target word (i.e., they did not differentially bias toward one of the two target words), and the same sentential contexts were recorded for both target words. The target words were then removed from these sentences and were replaced with the clear, noise, or silence versions from Experiment 1. In this way, the sentences were identical for all three types of target words. Although all of the target words had originally been produced in isolation, rather than in sentence context, the present sentences were spoken slowly enough that they did not appear out of place. The final stimuli were, “Look at that ____! Where is that ____? Look at the ____! Find the ____!”

Procedure. The procedure was identical to that in Experiment 1, except that each child heard only one version of the target words (clear, silence, or noise), rather than hearing all three versions. The toddlers received a total of 12 test trials, as in the earlier experiment. This resulted in six repetitions each of “kitty” and “doggie” in their assigned condition. (One child in the clear condition became fussy at the end of the experiment, and the last 2 trials, one of each word, were dropped.) Maintaining the total number of trials necessarily meant increasing the number of trials per condition; to ensure that boredom was not a factor, we examined the results in two ways: first, using all 12 trials, and second, using only the first 2 trials of each type.

Coding and analysis occurred as before. Reliability across coders for total looking was .95 for the clear condition, .96 for the noise condition, and .94 for the silence condition. Reliability on longest looks was .95 for both the clear and the noise conditions and .93 for the silence condition. Average differences were less than 0.5 sec in all cases.

Results and Discussion

The toddlers hearing the clear condition stimuli watched the matching video 2.03 sec longer than the mismatching video, a significant difference [$t(11) = 6.37, p < .0001$]. In addition, the toddlers watched the dog video longer on the “doggie” trials than on baseline trials [24.8%; $t(11) = 4.52, p < .0005$], although they did not watch the cat video longer on the “kitty” trials than on baseline trials [2%; $t(11) = 0.39, p > .10$]. The longest look measures showed stronger results, with longer looking toward the named video than toward that video in the baseline condition both for the cat [0.84 sec; $t(11) = 3.06, p < .01$] and for the dog [1.95 sec; $t(11) = 5.45, p < .0001$]. These results corroborated those from Experiment 2, showing that the children understood the meanings of both words and would watch the appropriate video when they recognized one of these words (see Figure 2).

We then analyzed the silence condition. The toddlers watched the matching video 1.56 sec longer than the mismatching video [$t(11) = 3.86, p < .005$]. This effect was present for both the cat [1.44 sec; $t(11) = 3.34, p < .005$] and the dog [1.35 sec; $t(11) = 3.06, p < .01$] in the longest looks but was significant only for the cat in the overall percentages of looking [11.7%; $t(11) = 1.85, p <$

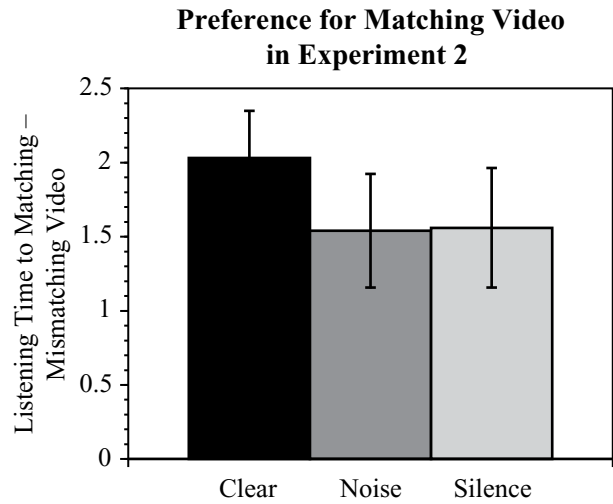


Figure 2. Differences between the time the subjects spent watching the matching video versus the mismatching video in Experiment 2.

.05]. For the dog, the effect was marginal [10.5%; $t(11) = 1.47, p < .10$]. As in Experiment 1, the toddlers could use partial information to help them determine the appropriate referent (as is shown in the far right column of Figure 2) but did not perform as well as with full information.

The effect was similar in the noise condition. Overall, the toddlers watched the matching video 1.54 sec longer than they did the mismatching video [$t(11) = 4.02, p < .001$]. This effect was significant in the longest looks [1.69 sec, $t(11) = 3.20$ for the cat; 1.01 sec, $t(11) = 3.37$ for the dog; both $ps < .005$]. It was significant only for the cat in the overall percentages of looking [14.6%, $t(11) = 1.81, p < .05$; for the dog, 8.25%, $t(11) = 1.22, p > .05$]. This is shown in the middle column of Figure 2. The complete set of data is shown in Table 2.

Overall, there was less difference between the three conditions in this experiment than in the prior one. Whereas all 12 toddlers looked longer at the appropriate video in

Table 2
Overall Looking Results in Experiment 2

	Condition					
	Clear		Silence		Noise	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Overall duration of looking (sec)						
To matching video	4.64	0.68	4.16	0.29	4.08	0.26
To mismatching video	2.62	0.51	2.60	0.19	2.54	0.17
Percentage of looking						
To cat, baseline trials	60.5	2.9	52.0	3.7	50.6	4.0
To dog, baseline trials	39.5	2.9	48.0	3.7	49.4	4.0
To cat, “kitty” trials	62.5	3.1	63.7	4.7	65.2	5.8
To dog, “doggie” trials	64.3	4.3	58.5	6.0	57.7	4.2
Longest looks (sec)						
To cat, baseline trials	2.60	0.19	2.29	0.35	2.16	0.20
To dog, baseline trials	1.83	0.20	1.74	0.11	2.10	0.26
To cat, “kitty” trials	3.43	0.24	3.74	0.37	3.85	0.39
To dog, “doggie” trials	3.79	0.36	3.09	0.44	3.11	0.27

the clear condition, 10 looked longer at the appropriate video in the noise condition, and 10 in the silence condition. The between-subjects analysis may have improved performance in both the noise and the silence conditions, but it did not appear to do so differentially. Looking at the statistics above, it is striking how comparable the performance was in the noise and the silence conditions. Comparing across these two conditions, we find no difference between the preference for the matching over the mismatching video in the two conditions [$t(22) = -0.03$, $p > .05$]. Moreover, neither the toddlers' total looking nor the toddlers' longest looking differed across conditions (both F s < 1). Even in a between-subjects analysis, there was no evidence that the toddlers restored the missing phoneme.

As was mentioned earlier, it is possible that hearing six trials of each type was somewhat boring to the young toddlers. We therefore reexamined the data, looking only at the first two trials of each type (the same number of trials per type as that examined in Experiment 1). The results for the clear condition are nearly identical when one looks only at the first two trials of each condition as when one looks at all six trials. The toddlers watched the matching video 2.37 sec longer than the mismatching video, a significant difference [$t(11) = 5.41$, $p < .0005$]. They watched the dog video longer on the "doggie" trials than on baseline trials [22.3%; $t(11) = 3.47$, $p < .005$], although they did not watch the cat video longer on the "kitty" trials than on baseline trials [8.3%; $t(11) = 1.13$, $p > .10$]. As in the earlier analyses, the longest look measures appear to be more sensitive; according to this measure, the children watched the dog video longer on "doggie" trials [1.61 sec; $t(11) = 3.67$, $p < .005$] and the cat video longer on "kitty" trials [1.35 sec; $t(11) = 2.83$, $p < .01$].

The results for the silence condition are somewhat weaker when one looks only at the first two trials. Children watched the matching video 1.31 sec longer than the mismatching video [$t(11) = 3.15$, $p < .005$]. But although this effect was present for both the cat [1.51 sec; $t(11) = 3.31$, $p < .005$] and the dog [1.25 sec; $t(11) = 2.07$, $p < .05$] in the longest looks, it was not significant in either one on the basis of the overall percentages of looking [for the cat, 11.2%, $t(11) = 1.59$, $p < .10$; for the dog, 6.3%, $t(11) = 0.72$, $p > .10$].

The same weakening of the overall effect was found in the noise condition. Overall, the children watched the matching video 1.43 sec longer than the mismatching video [$t(11) = 2.40$, $p < .05$]. The advantage for the matching video was not different for the children in the noise condition than for those in the silence condition [$t(22) = 0.17$, $p > .05$]. Looking at the two videos separately, the effect was significant for the cat in both the longest looks and the overall percentages of looking [1.98 sec, $t(11) = 2.71$, $p < .05$, and 20.6%, $t(11) = 2.21$, $p < .05$, respectively] but was not significant for the dog in either case [0.52 sec, $t(11) = 1.13$, and 0.5%, $t(11) = 0.06$; both p s $> .05$]. Clearly, then, the increase in the number of trials per condition did not reduce the overall effect of restoration; if anything, the effects for *both* the

noise and the silence conditions were somewhat weaker when only two trials were looked at.

In general, the difference between the three conditions in the present experiment was smaller than that in Experiment 1. This may have been a result of the lack of cross-trial comparisons, or it may have been the result of the increase in the number of trials per condition. Moreover, although there was a significant improvement in the silence condition over the noise condition in Experiment 1, the difference between the two here was negligible. But the general conclusion from this experiment is quite similar to that from Experiment 1: There is no evidence that the children experienced any increase in intelligibility from having noise fill the gaps in the signal.

EXPERIMENT 3

The toddlers in the previous two experiments failed to demonstrate evidence of perceptual restoration. However, the lack of perceptual restoration may have been a result of the particular words that were presented. Although restoration tends to be strong with both stop consonants and fricatives in adult listeners (Samuel, 1981b), this fact may not carry over to children. Moreover, stop consonants may be more likely to be heard when replaced by silence than are other sounds, at least word internally. Restoration also tends to be stronger with longer words (Samuel, 1981a). In short, perhaps we would have found different results with the children if we had used different words.

A better test would involve words that both were much longer and consistently resulted in the illusion of restoration for adults. Unfortunately, we are somewhat limited in the stimuli we can present to young children. In order to test for restoration, both words need to be well known to the toddlers, which is less likely to be the case for longer words. The words also need to be clearly and distinctly picturable, and the images need to be of roughly similar visual interest. (If one item was more interesting visually than the other, the toddlers might choose to focus solely on that object, preventing us from testing the toddlers' performance on the alternative object.)

Since stop consonants were used in our previous two experiments, we decided to search for words containing a fricative for the present experiment. We were unable to come up with alternative items that would be known by children at only 24 months of age. Thus, the present experiment focused on children slightly older than those in the prior experiments (30 months of age). The two words selected were "dinosaur" and "lobster." Both words are slightly longer than those used in the prior experiments (thus providing more contextual information as to the words' identity) and contain a medial /s/, which may be a better phoneme for testing restoration than are stop consonants. The fricative is part of a cluster in one of the two words, which may cause the two words to differ slightly, but there is no evidence in the adult literature that this should affect restoration. Although the word "lobster" may be less well-known than "dinosaur" to children of this age, our initial conversations with parents suggested

that children of this age did know this word: Tanks of live lobsters are a common sight in area grocery stores.

In addition, we performed a more rigorous set of adult pilot testing in the present experiment, and our final test items were restored 100% of the time in noise and 0% of the time in silence. That is, adult listeners always heard the absent /s/ in the noise condition and never reported hearing it in the silence condition. Although increases in word intelligibility when gaps are filled with noise often occur even without the illusion of restoration, finding stimuli that consistently produce this illusion should increase the likelihood of finding an intelligibility difference in children. These items therefore provide a stronger test of perceptual restoration in children.

Method

Subjects. Twenty-three 30-month-old children (9 of them female and 14 male; mean age, 30.1 months; range, 29.1–31.1) participated in this experiment. Data from 5 additional children were dropped from analysis for equipment failure ($n = 1$), fussiness ($n = 2$), or experimenter error ($n = 1$) or for not having English as the primary language. Finally, 1 child wore glasses, which picked up a glare from the television sets; this made her looking behavior very difficult to code, and her data were therefore not included.

Stimuli. The visual stimuli were videotapes of two small stuffed animals (a dinosaur and a lobster) against a white background. We felt static pictures would be somewhat uninteresting to our subjects and, therefore, videotaped the two toys “dancing” by means of a fishing line attached to the experimenter’s hand (which was out of view of the camera). Each video lasted approximately 10 min (the actual test sessions took between 5 and 8 min, so the last few minutes of tape were never seen).

The auditory stimuli were recorded in the same manner and by the same speaker as in Experiment 1. All the items were recorded in a noise-reducing chamber at a 44.1-kHz sampling rate, 16-bit resolution, and were stored on computer disk. The speaker first recorded a number of instances of the target words. In order to ensure that there were no transitional cues for the missing phoneme, all the target words for these items were mispronounced prior to recording. Cut-points were selected on the basis of visual inspection and a spectral analysis of the recordings, so as to minimize any transitional cues to the medial fricative. This fricative was removed and replaced with either noise or silence. For this experiment, the noise consisted of a mix of white noise and a 400-Hz tone. In addition, a clear version of each word was recorded. Initial pilot testing with adults suggested that the restoration illusion was stronger with lower sampling rates. We therefore downsampled each stimulus to 16 kHz and filtered at the Nyquist frequency, so that information higher than 8 kHz was not present in the signal.

A number of tokens of both the noise and the silence versions of each word were presented to 10 adult subjects who were unaware of the purposes of the experiment. These listeners were asked to write down what they thought they heard on each trial. The single best version of each word was selected on the basis of these reports. These tokens were identified as containing the missing /s/ in the noise condition by all 10 listeners and were never identified as containing the /s/ in the silence condition. (Typical responses for the silence condition were “dinah ore” and “lop ter,” although spellings varied among the listeners.)

The speaker next recorded a series of sentences regarding each target item (“Look at the lobster!” “Where is that dinosaur?” etc.). The target words were removed, and these served as frames for all three conditions. The best exemplar of each word was then inserted in these frames, creating sets of sentences with the identical target word. This method of editing was also used with a single clear condition item, so that the clear versions of the sentences contained the

same lack of variability among target word tokens as that found in the noise and silence conditions.

Procedure. The results of Experiment 2 suggest that even without the possibility of cross-trial comparisons, the toddlers still did not perform better with noise than with silence. Moreover, responses to the clear condition seemed stronger in Experiment 1 (using the within-subjects design and fewer trials per condition) than in Experiment 2 (using the between-subjects design and more trials per condition). Finally, given the increased difficulty of the words themselves, we wanted to ensure that all the children actually knew both words, and this required that each child be tested with clear condition stimuli. For these reasons, we used the same procedure here as in Experiment 1. Practice trials (with only a single video) again lasted 5 sec; test trials and baseline trials lasted 13 sec (the length of the auditory stimuli).

Coding and analysis occurred as before. Correlations across coders were .96 for the overall durations and .95 for single longest looks; on average, the coders differed by less than 0.6 sec per trial for both longest and total looking.

Results and Discussion

In the clear condition, the children watched the matching video 2.93 sec longer than the mismatching video, a significant difference [$t(22) = 4.00, p < .0005$]. In addition, the toddlers watched the dinosaur video longer on the “dinosaur” trials than on baseline trials [14.7%; $t(22) = 2.57, p < .01$] and watched the lobster video longer on the “lobster” trials than on baseline trials [12.1%; $t(22) = 2.19, p < .05$]. The longest look measures showed similar results, with longer looking toward the named video than toward that video in the baseline condition for both the lobster [2.98 sec; $t(22) = 4.57, p < .0005$] and the dinosaur [1.70 sec; $t(22) = 2.89, p < .005$]. These results show that the children understood the meanings of both words. The advantage for the dinosaur was slightly larger than that for the lobster on the overall looking measure but was slightly smaller on the longest look measure. Neither difference was significant, however [for overall proportions, $t(22) = 0.30$; for longest looks, $t(22) = 1.32$; both $ps > .05$]; there does not appear to be any evidence that the children’s knowledge of the word “lobster” was different from that for the word “dinosaur.”

We then analyzed the silence condition. The children watched the matching video 3.4 sec longer than the mismatching video [$t(22) = 5.53, p < .0001$]. This effect was present for both the lobster [14.8%; $t(22) = 2.94, p < .005$] and the dinosaur [15.5%; $t(22) = 2.72, p < .01$] and was similarly present in the longest look data [3.06 sec for the lobster, $t(22) = 3.72, p < .001$; 2.28 sec for the dinosaur, $t(22) = 4.32, p < .0005$]. This can be seen in Figure 3, which shows the time the subjects spent watching the matching and mismatching videos for this experiment; Table 3 gives the complete set of data.

The effect was also present in the noise condition. Overall, the children watched the matching video 2.06 sec longer than they did the mismatching video [$t(22) = 3.83, p < .0005$]. This effect was significant for the dinosaur video [13.7%; $t(22) = 2.24, p < .05$], but not for the lobster video [7.0%; $t(22) = 1.44, p < .10$]. However, the single longest looks were significant for both videos [1.91 sec for the lobster, $t(22) = 3.87, p < .0005$; 1.68 sec for the dinosaur, $t(22) = 3.03, p < .005$].

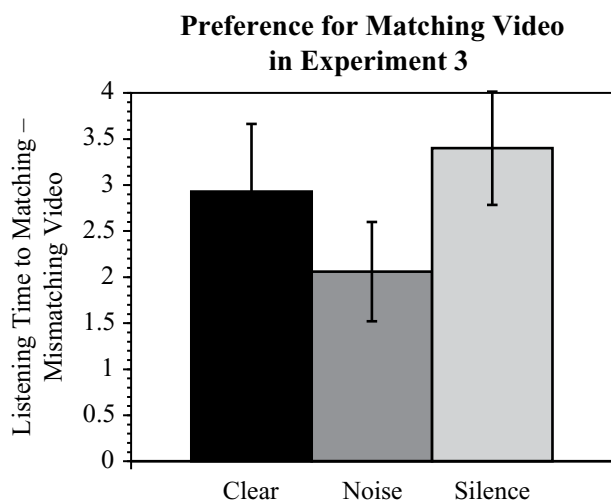


Figure 3. Differences between the time the subjects spent watching the matching video versus the mismatching video in Experiment 3.

Comparing across the noise and silence conditions, we find that the effect was significantly larger for silence than for noise. For the time the subjects spent watching the matching versus mismatching video, there was a significant effect of match [$F(1,22) = 31.15, p < .0001$] but also a significant interaction with type [$F(1,22) = 4.70, p < .05$]. The children's preference for the matching video was 3.4 sec for the silence condition but only 2.1 sec for the noise condition. Looking at the percentages of looking and the longest looking data shows a similar pattern, but one that just misses significance. Averaging across the two animals, the two conditions differ only marginally in percentage of looking [15.2% above baseline in the silence, but only 10.4% in the noise; $t(22) = 1.76, p < .10$] and for longest looks [2.67 sec for silence, 1.80 sec for noise; $t(22) = 1.80, p < .10$]. However, the trend occurred for

both videos. For the lobster video, the children's percentages of looking were 7% longer, and their single longest looks were over 1 sec longer, in the silence condition. For the dinosaur video, the children's percentages of looking were 1.5% longer and their single longest looks over a half second longer in the silence condition than in the noise condition. Clearly, the toddlers did not perform significantly better in the noise condition than in the silence condition. Indeed, the significant interaction based on matching data suggests that the children were significantly worse in noise than in silence, rather than better (as was the case in Experiment 1). This suggests that the lack of an advantage for the noise stimuli in these experiments is not simply the result of a lack of power. Instead, these comparisons suggest that while children can use their prior knowledge of words to interpret partial information correctly, they do not appear to experience an increase in intelligibility when gaps in the signal are filled with noise.

Across the three conditions, 18 of the 23 toddlers looked longer at the appropriate (as compared with the inappropriate) video in the clear condition, and 20 did so in the silence condition. Yet only 15 did so in the noise condition. Although the silence condition may have reached ceiling performance, the noise condition clearly did not.

One additional concern, however, is that the trial length in this experiment (and in the previous two experiments) may have been overly long. Perhaps the children were looking appropriately at the beginning of trials but then began to look randomly.⁶ To examine this, we reanalyzed the data from all three experiments, looking only at the initial portion of each trial. Looking behavior was analyzed for the initial 3 sec of each trial (i.e., from the point at which the first token of the target word was completed until a point 3 sec later, regardless of where in a sentence this occurred).

For the results from Experiment 1, this reanalysis did not lead to any increased evidence of restoration. There was an overall effect of matching [$F(1,19) = 15.63, p < .001$]

Table 3
Overall Looking Results in Experiment 3

	Condition					
	Clear		Silence		Noise	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Overall duration of looking (sec)						
To matching video	6.74	0.37	7.19	0.34	6.28	0.32
To mismatching video	3.82	0.42	3.79	0.34	4.21	0.32
Percentage of looking						
To lobster, "lobster" trials	73.2	4.4	75.9	4.7	68.1	4.7
To dinosaur, "dinosaur" trials	53.7	4.8	54.4	5.2	52.6	5.0
Longest looks (sec)						
To lobster, "lobster" trials	6.52	0.54	6.60	0.72	5.45	0.58
To dinosaur, "dinosaur" trials	4.08	0.53	4.66	0.51	4.06	0.49

Note—In all the statistics on percentages of looking and longest looks, the difference scores between the numbers above and the values obtained on baseline trials were examined. That is, what was crucial was not the number of seconds the lobster was watched on "lobster" trials but whether there was an increase above that found on baseline trials. On baseline trials, overall percentages of looking were 61.1% to the lobster ($SE = 2.5\%$) and 38.9% to the dinosaur ($SE = 2.5\%$). Longest looks were 3.54 sec ($SE = 0.33$) to the lobster and 2.38 sec ($SE = 0.16$) to the dinosaur.

and a significant interaction with condition [$F(2,38) = 6.33, p < .005$]. The effect of matching was significant in the clear condition [$t(19) = 5.42, p < .0001$] and in the silence condition [$t(19) = 1.95, p < .05$], but not in the noise condition [$t(19) = 1.46, p < .10$]. The size of the effect was larger in the clear condition than in either of the other two conditions, which did not differ from one another [clear vs. noise, $t(19) = 4.71, p < .0005$; clear vs. silence, $t(19) = 2.07, p < .05$; noise vs. silence, $t(19) = 0.94, p > .05$]. This suggests that although the children knew the words, they did not perform differently in the noise condition than in the silence condition. Results from the two videos separately show similar results. On the basis of longest look measures, the children looked longer at both the cat and the dog in the clear condition than in the baseline conditions [cat, $t(19) = 3.33, p < .005$; dog, $t(19) = 4.43, p < .0005$] but did not do so for the noise condition [cat, $t(19) = 1.41, p < .10$; dog, $t(19) = 1.72, p < .10$] and did so only for the dog in the silence condition [cat, $t(19) = 1.51, p < .10$; dog, $t(19) = 2.59, p < .01$]. The percentage of looking results gave less clear results. Although the children did not look longer to the target item in either the noise or the silence condition than in the baseline condition [cat in the noise condition, $t(19) = -0.79, p > .05$; dog in the noise condition, $t(19) = 0.96, p > .05$; cat in the silence condition, $t(19) = -0.27, p > .05$; dog in the silence condition, $t(19) = 1.21, p > .05$], they did so for the clear condition only for the dog, not for the cat, making these results generally difficult to interpret [cat, $t(19) = 1.57, p < .10$; dog, $t(19) = 3.80, p < .001$]. Clearly, though, looking only at the initial portion of each trial does not provide any evidence for improved performance in the noise condition over that in the silence condition. Nor were the children at ceiling performance, since they consistently performed better in the clear condition than in either of the other two conditions.

The reanalysis of the results from Experiment 3 likewise showed no benefit of the noise and, in fact, still showed that the silence condition led to significantly better performance than did the noise condition. There was again an overall effect of matching [$F(1,21) = 21.45, p < .0001$] and a significant interaction with condition [$F(2,42) = 4.52, p < .05$]. The effect of matching was significant in the clear condition [$t(21) = 3.42, p < .005$] and in the silence condition [$t(21) = 5.89, p < .0001$], but not in the noise condition [$t(21) = 1.09, p > .05$]. Moreover, the size of the effect was similar in the clear and the silence conditions, both of which differed from the noise condition [clear vs. noise, $t(21) = 2.11, p < .05$; silence vs. noise, $t(21) = 2.75, p < .01$; but clear vs. silence, $t(21) = -0.75, p > .05$]. Results from both the longest look measures and the percentage of looking measures were similar. In the clear condition, the children looked longer at both the lobster [longest look, $t(21) = 3.03, p < .005$; percentage of looking, $t(21) = 2.69, p < .01$] and the dinosaur [longest look, $t(21) = 1.95, p < .05$; percentage of looking, $t(21) = 1.96, p < .05$]. In the silence condition, the children also looked longer at the appropriate videos [for the lobster, longest look, $t(21) = 3.91, p < .0005$,

and percentage of looking, $t(21) = 4.24, p < .0005$; for the dinosaur, longest look, $t(21) = 2.98, p < .005$, and percentage of looking, $t(21) = 1.94, p < .05$]. But in the noise condition, the children never looked longer at the appropriate video [for the lobster, longest look, $t(21) = 1.71, p < .10$, and percentage of looking, $t(21) = 1.13, p > .05$; for the dinosaur, longest look, $t(21) = 0.64, p > .05$, and percentage of looking, $t(21) = 0.61, p > .05$]. The results from this experiment show quite clearly that the children performed *worse* in the noise condition, not better—and thus, that lack of power, or the performance variability inherent in testing infants, is not likely to be the reason for the lack of a noise advantage.

The results from Experiment 2 are somewhat more ambiguous as a result of ceiling performance. Here, there was an overall effect of matching [$F(1,33) = 42.48, p < .0001$] but no interaction with type. The effect of match was significant in all three conditions [clear, $t(10) = 4.36, p < .001$; noise, $t(10) = 4.26, p < .001$; silence, $t(10) = 2.62, p < .05$]. This suggests that the children may have reached ceiling performance in the noise and silence conditions. However, the effect was driven only by the “doggie” trials, not by the “kitty” trials. For the clear condition, the children did not look longer to the cat on “kitty” trials by either measure [longest look, $t(10) = 0.51$; percentage of looking, $t(10) = 0.41$; both $ps > .05$], making interpretation of those trials difficult. On dog trials, the children did look appropriately in the clear condition [longest look, $t(10) = 3.90, p < .005$; percentage of looking, $t(10) = 3.09, p > .01$] but did so only by one of the two measures in both the noise and the silence conditions [noise condition, longest look, $t(10) = 1.12, p > .05$, and percentage of looking, $t(10) = 2.18, p < .05$; silence condition, longest look, $t(10) = 1.98, p < .05$, and percentage of looking, $t(10) = 1.17, p > .05$]. Thus, there is again no evidence for improved performance in the noise condition than in the silence condition. The lack of an interaction in the matching results suggest that this may, in part, have been the result of ceiling performance in both the noise and the silence conditions. However, the results from the individual trials suggest that the performance in the dog trials was not as good in either the noise or the silence condition as it was in the clear condition; performance on cat trials seems to have been poor even in the clear condition. Apparently the “ceiling” performance here was the result of poor performance in the clear condition, rather than good performance in the other two conditions. This poorer clear condition performance in the reanalysis may be an indication that the longer trial lengths were necessary to adequately capture the children’s true knowledge.

Looking across these three experiments, the reanalysis does not appear to provide any additional evidence for restoration. If anything, the reanalysis of Experiment 3 suggests even more strongly that performance was worse when a phoneme was replaced by noise than when it was replaced by silence.

One final concern rests on the nature of the stimuli themselves. All of the items selected for restoration in these experiments had been misproduced prior to edit-

ing. That is, the speaker had produced sequences such as “dinoshore,” rather than “dinosaur,” before replacing the fricative with silence or noise. This approach has been used frequently in the adult literature and ensures that any transitional cues remaining after editing will indicate an incorrect phoneme and, thus, will be unlikely to lead to restoration. However, young toddlers may be more open to new lexical entries than are adults; perhaps they *did* show perceptual restoration but restored the items to their original productions, rather than to the target words.⁸ That is, perhaps the children did restore the items but did so to a “new” lexical entry, “dinoshore,” rather than to the intended “dinosaur.” This explanation is of particular concern given that the children performed more accurately in silence than in noise in Experiment 3; this is what would be expected if noise led to the perception of a different word (or of a conflicting phoneme), whereas silence simply involved a missing phoneme. To examine this, Experiment 4 replicated Experiment 3, but the original, accurate productions of “lobster” and “dinosaur” were used as the basis for creating the noise and silence conditions. Furthermore, given the somewhat varying results from the within-subjects analyses in Experiments 1 and 3 and the between-subjects analysis in Experiment 2, both approaches were used in Experiment 4.

EXPERIMENT 4

The present experiment was an attempt to replicate the findings from the earlier experiments, using natural (accurate) productions as the basis for the noise and silent versions of the stimuli. If the results from this experiment mirrored those of the prior experiments, it would support the notion that young children do not show perceptual restoration. If children showed evidence of restoration in this experiment, it would suggest that the lack of an advantage for noise stimuli over silence stimuli in the earlier experiments was the result of the particular method of creating the stimuli (i.e., misproducing the target phoneme prior to editing).

Method

Subjects. Fifty-six 30-month-old children (28 of them female and 28 male; mean age, 30.3 months; range, 28.8–31.8) participated in this experiment. Data from 25 additional children were dropped from analysis for fussiness ($n = 9$), equipment failure ($n = 3$), or experimenter error ($n = 8$), for not having English as the primary language ($n = 1$), or for missing data ($n = 4$). In addition, 5 mothers either lost their visor or peeked; the data from these toddlers were also excluded. Of the final subjects, 20 were assigned to the within-subjects condition (12 of them male and 8 female; mean age, 30.7 months), and 36 were assigned to the between-subjects condition (16 of them male and 20 female; mean age, 30.2). Of the latter, the average age for those in the clear condition was 30.1 months (range, 28.8–31.0), for those in the noise condition, 30.2 months (range, 29.5–30.9), and for those in the silence condition, 30.4 months (range, 29.7–31.3); these ages did not differ from one another statistically.

Stimuli. The visual stimuli were identical to those in Experiment 3. The clear auditory stimuli were also identical to those in the prior experiment. New versions of the noise and silence versions were created, using the same methods as before, but using the clear

versions as the base stimuli. This resulted in target items that were acoustically identical in all respects, except for the target phoneme (which was present for the clear stimuli and was replaced with either silence or noise in the other conditions).

Procedure. Procedure, coding, and analysis occurred as before. The children participated in either a within-subjects condition [12 test trials, 2 of each combination of animal (lobster vs. dinosaur) and condition (noise, silence, and clear)] or a between-subjects condition (12 test trials, 6 of each animal, with each child hearing only one type of condition). Correlations across coders were .90 for the overall durations and .84 for single longest looks in the within-subjects condition. Average difference scores were less than 0.7 sec in both conditions. For the between-subjects condition, correlations for the longest looks were .94 in the clear condition and .96 in both the noise and the silence conditions; for total looking, .96 in both the clear and the noise conditions and .97 in the silence condition. Average difference scores were less than 0.5 sec in all the conditions.

Results and Discussion

We first examined the data from the between-subjects condition. Toddlers in this condition heard one of three types of stimuli: the clear stimuli, the noise stimuli, or the silent stimuli. The toddlers hearing the clear stimuli watched the matching video 2.33 sec longer than the mismatching video, a significant difference [$t(11) = 4.30, p < .001$]. In addition, the toddlers watched the dinosaur video longer on the “dinosaur” trials than on baseline trials [15.8%; $t(11) = 4.28, p < .001$], although they did not watch the lobster video longer on the “lobster” trials than on baseline trials [6.7%; $t(11) = 1.03, p > .10$]. The longest look measures showed stronger results, with longer looking toward the named video than toward that video in the baseline condition for both the lobster [2.33 sec; $t(11) = 3.08, p < .01$] and the dinosaur [2.05 sec; $t(11) = 4.38, p < .001$]. As in the prior experiments, these results suggest that the children understood the meanings of both words and would watch the appropriate video when they

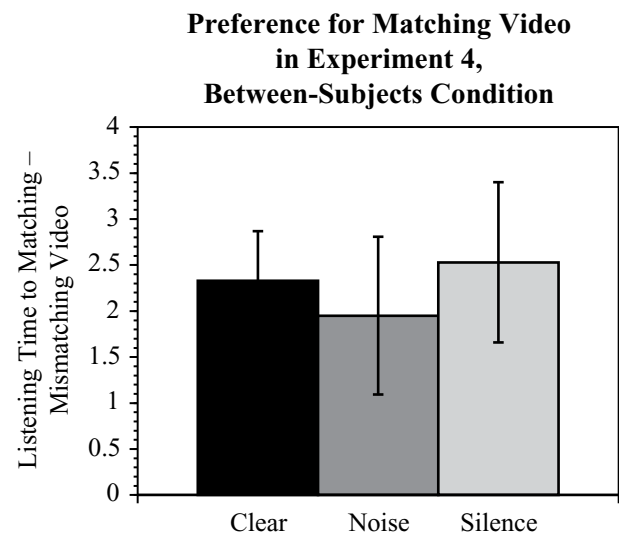


Figure 4. Differences between the time the subjects spent watching the matching video versus the mismatching video in Experiment 4, between-subjects condition.

recognized one of these words (see Figure 4, left-hand column).

We then analyzed the silence condition and found the same pattern of results as in the clear condition. The children watched the matching video 2.54 sec longer than the mismatching video [$t(11) = 2.91, p < .01$]. This effect was present for both the lobster [2.44 sec; $t(11) = 3.12, p < .005$] and the dinosaur [2.03 sec; $t(11) = 5.67, p < .0001$] in the longest looks but was significant only for the dinosaur for the overall percentages of looking [11.8%; $t(11) = 2.05, p < .05$], not for the lobster [10.5%; $t(11) = 1.26, p > .05$]. This can be shown in the far right column of Figure 4.

The noise condition showed a similar pattern, except that the percentages of looking data were not significant for either animal. Overall, the children watched the matching video 1.99 sec longer than they did the mismatching video [$t(11) = 2.35, p < .05$]. This effect was significant in the longest looks [2.92 sec, $t(11) = 3.24, p < .005$ for the lobster and 1.61 sec, $t(11) = 2.21, p < .05$ for the dinosaur], but not in the overall percentages of looking [12.2%; $t(11) = 1.70, p < .10$, for the lobster; 6.8%, $t(11) = 0.95, p > .05$, for the dinosaur]. This can be shown in the middle column of Figure 4. The complete set of data is shown in Table 4.

Comparing across the three conditions, 10 of the 12 toddlers looked longer at the appropriate video in the clear condition, 10 of 12 did so in the silence condition, and

9 of 12 did so in the noise condition. Indeed, looking at the matching data for the noise and the silence conditions shows a significant effect of match [$F(1,22) = 13.87, p < .001$] but no interaction with stimulus condition ($F < 1$). Comparing across all three conditions likewise led to no interactions. Apparently, the preference for the matching video is similar in size for the clear, silence, and noise conditions [$t(22) = 0.46$ for silence vs. noise, $t(22) = 0.21$ for clear vs. silence, and $t(22) = 0.34$ for clear vs. noise; all $ps > .05$]. Moreover, neither toddlers' total looking nor toddlers' longest looking differed across conditions (both $Fs < 1$). There is still no evidence that the toddlers restored the missing phoneme; indeed, the same trend occurred as in the prior experiments, with better performance for the silence condition than for the noise condition. Yet, as in the other between-subjects condition, there is also a concern that the children may have been at ceiling performance, since the children watched the matching video as long in the silence condition as in the clear condition.

The results from the within-subjects condition look similar, although there appears to be a greater difference between the clear condition and the other two conditions here (see Figure 5). All but 1 of the children looked longer at the appropriate video in the clear condition, 19 out of 20 did so in the silence condition, and 16 out of 20 did so in the noise condition. The children looked longer at the matching than at the mismatching video in all three

Table 4
Overall Looking Results in Experiment 4

	Condition					
	Clear		Silence		Noise	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Within Subjects						
Overall duration of looking (sec)						
To matching video	7.35	0.36	6.74	0.46	6.45	0.43
To mismatching video	3.12	0.30	3.30	0.29	3.59	0.29
Percentage of looking						
To lobster, "lobster" trials	74.0	3.6	66.2	4.0	64.8	4.1
To dinosaur, "dinosaur" trials	64.9	3.3	67.8	3.8	63.6	4.5
Longest looks (sec)						
To lobster, "lobster" trials	5.48	0.54	5.11	0.50	4.79	0.46
To dinosaur, "dinosaur" trials	5.18	0.45	5.20	0.56	5.30	0.60
Lobster baseline trials: 56.4% (<i>SE</i> = 2.3%) for percentage of looking, 3.41 sec (<i>SE</i> = 0.32) for longest looks						
Dinosaur baseline trials: 43.6% (<i>SE</i> = 2.3%) for percentage of looking, 2.29 sec (<i>SE</i> = 0.16) for longest looks						
Between Subjects						
Overall duration of looking (sec)						
To matching video	6.25	0.29	6.27	0.50	6.36	0.48
To mismatching video	3.92	0.41	3.73	0.49	4.38	0.43
Percentage of looking						
To lobster, baseline trials	54.2	2.5	62.2	3.7	58.3	3.0
To dinosaur, baseline trials	45.8	2.5	38.3	3.7	41.8	3.0
To lobster, "lobster" trials	60.8	5.7	72.8	6.7	70.4	5.2
To dinosaur, "dinosaur" trials	61.6	3.3	50.1	4.6	48.4	6.6
Longest looks (sec)						
To lobster, baseline trials	2.93	0.29	2.89	0.23	2.85	0.30
To dinosaur, baseline trials	2.57	0.27	1.61	0.15	2.22	0.27
To lobster, "lobster" trials	5.26	0.70	5.34	0.71	5.77	0.82
To dinosaur, "dinosaur" trials	4.62	0.37	3.63	0.36	3.80	0.69

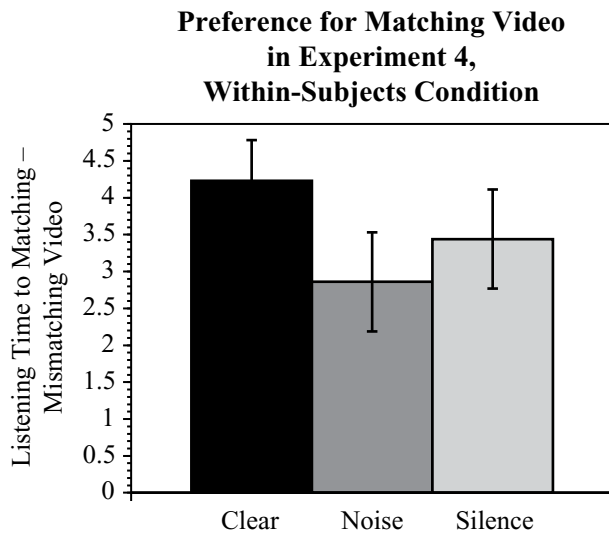


Figure 5. Differences between the time the subjects spent watching the matching video versus the mismatching video in Experiment 4, within-subjects condition.

conditions [clear, 4.23 sec, $t(19) = 7.66, p < .0001$; silence, 3.44 sec, $t(19) = 5.12, p < .0001$; noise, 2.86 sec, $t(19) = 4.25, p < .0005$]. Although the looking times toward the matching video are longest in the clear condition and shortest in the noise condition, an overall ANOVA across the three conditions shows an effect of matching [$F(1,19) = 54.46, p < .0001$] but only a trend toward an interaction with condition [$F(2,38) = 1.78, p > .05$]. That said, paired t tests do show a significant difference between the size of the matching effect in the clear versus the noise conditions [$t(19) = 2.16, p < .05$], although the silence condition does not differ from either [clear vs. silence, $t(19) = 0.92$; silence vs. noise, $t(19) = 0.87$; both $ps > .05$].

On the basis of longest looking times, the children looked at the appropriate video longer on named trials than on baseline trials in all six conditions [clear, 2.38 sec for lobster, 2.64 sec for dinosaur, $t(19) = 3.17, p < .005$, and $t(19) = 5.40, p < .0001$, respectively; for silence, 1.70 sec for lobster and 2.40 sec for dinosaur, $t(19) = 3.34, p < .005$, and $t(19) = 4.20, p < .0005$; for noise, 1.23 and 3.10 sec, $t(19) = 2.18, p < .05$, and $t(19) = 4.44, p < .0005$]. On average, the three conditions did not differ from one another [clear vs. noise, $t(19) = 1.05$; clear vs. silence, $t(19) = 1.08$; noise vs. silence, $t(19) = 0.32$; all $ps > .05$].

On the basis of overall percentages, the children looked at the appropriate video longer on named trials than on baseline trials in five of the six conditions [clear, 20.25% for lobster, 20.15% for dinosaur, $t(19) = 3.87, p < .001$, and $t(19) = 5.93, p < .0001$, respectively; silence, 10.65% for lobster and 22.40% for dinosaur, $t(19) = 2.78, p < .01$, and $t(19) = 4.58, p < .0005$; noise, 7.1% and 17.95%, $t(19) = 1.32, p > .05$, and $t(19) = 2.98, p < .005$]. Looking across the averages for the three condi-

tions, there is a marginal effect of type [$F(2,38) = 2.82, p < .10$]. Follow-up t tests show that the clear condition (20.2%) did differ significantly from the noise condition (12.53%) [$t(19) = 2.55, p < .01$], but the silence condition (16.5%) had an intermediate value that did not differ significantly from either [clear vs. silence, $t(19) = 1.03, p > .05$; silence vs. noise, $t(19) = 1.29, p > .05$].

Thus, on the basis of all three types of analyses, the children showed a trend toward best performance in the clear condition and worst in the noise condition, but the interactions across conditions were not statistically significant. What does seem clear, however, is that there is no evidence for perceptual restoration: Even when correctly produced items were used as the basis for stimulus creation, the children showed no evidence of enhanced performance in the noise condition, in comparison with the silence condition; indeed, all trends point in the opposite direction. Similarly, when looking at both the within- and the between-subjects variants, there is at no point any hint of an advantage for the noise condition over the silence condition.

As a final analysis, we performed an outright comparison between the within- and the between-subjects designs, to determine whether the methodology used led to significant differences. We first compared performance in the clear condition. Here, we found significantly longer looking for the matching video across both designs [$F(1,30) = 76.86, p < .0001$]. But although there was no overall effect of the methodology ($F < 1$), there was an interaction [$F(1,30) = 5.28, p < .05$]. The children in the within-subjects design showed a greater preference for the matching video than did the children in the between-subjects design; that is, those children who heard 12 clear trials did not perform as well as the children who heard only 2 clear trials, suggesting either that the clear trials were more obvious when the noise and silence trials were present (in the within-subjects condition), or that the children in the between-subjects condition were simply getting bored after hearing the identical stimuli more often.

This did not appear to be the case for the other two conditions. For the silence condition, there was an overall effect of match [$F(1,30) = 33.98, p < .0001$] but no overall effect of design and no interaction between the two factors (both $Fs < 1$). Similarly, for the noise condition, there was an overall effect of match [$F(1,30) = 23.11, p < .0001$] but no overall effect of design [$F(1,30) = 2.40, p > .10$] and no interaction between the two factors ($F < 1$). Thus, it appears that the methodology changed performance on the clear trials, but not on the noise or silence trials. If cross-trial comparisons occurred, one might expect that the noise and silence trials would seem "less good" in comparison with the clear trials and, thus, that the interactions would occur there. Instead, this pattern seems to suggest that the children got bored with the clear trials when there were large numbers of them and, thus, that the clear condition is not a good measure of "ceiling" performance in the between-subjects analysis. But why did the children show this effect of boredom only for the clear trials, and not for the other trials? Much of the infant literature suggests that

children get bored more quickly with “simpler” tasks, preferring to look at familiar items when processing is more difficult and looking away from these items when processing is more advanced (Hunter & Ames, 1988; Rose, Gottfried, Melloy-Carminar, & Bridger, 1982). Perhaps these toddlers show the same pattern, and looking appropriately to a clear exemplar was simply a much easier task for them than looking appropriately to a less ideal exemplar. Although this interpretation is quite speculative, it would support the general conclusion that toddlers need more of an acoustic signal to be intact in order to easily identify it; looking appropriately to an item with a phoneme missing would appear to be a difficult task for them. Regardless of the explanation, the findings do suggest that the effect of methodology (between vs. within) did not influence how the children performed with the noise and silence stimuli but influenced performance only with the clear items.

GENERAL DISCUSSION

The present experiments demonstrate that children can use prior knowledge of the language to help them interpret partial information. Given only part of a word, toddlers appear able to determine which of two referents that word refers to (see Swingley, Pinto, & Fernald, 1999, for similar findings). In all but one analysis (that of the noise condition in Experiment 1), the children looked longer at the appropriate (matching) video in the silence and noise conditions than at the inappropriate video. This suggests that the children were able to identify which object better matched the incomplete audio signal they heard.

However, young children appear in general to need more of the signal to be intact in order to properly identify the stimulus. Even with only two stimulus choices in front of them, the children generally did not identify the words as accurately when a phoneme was missing as when the entire sequence was present. For example, in Experiment 1, 19 out of 20 children showed longer looking to the appropriate video in the clear condition; only 12 showed longer looking to the appropriate video in the silence condition, and only 10 in the noise condition. Thus, children do not seem to be able to fully correct for the missing information, even with strong environmental cues to word meaning (i.e., the two videos). Even in situations in which children perform quite well in good listening conditions, they appear to show marked decrements relative to adults in more difficult listening conditions.

This requirement for additional acoustic input in order to interpret items correctly has been demonstrated in other types of tasks as well. For example, young children apparently require larger acoustic differences to discriminate sounds than do older children or adults (Elliott, Busse, Partridge, Rupert, & DeGraaff, 1986; Elliott, Hammer, Scholl, & Wasowicz, 1989). They also require larger amplitude differences (Elliott, Clifton, & Servi, 1983) and larger signal-to-noise ratios (Elliott et al., 1979) in order to identify individual words. The methodologies used in these prior studies were very different from that in the present

study, and the researchers focused on different types of information (size of acoustic difference, rather than proportion of the word presented). However, the similarities in findings suggest that all of these may be examples of a more global phenomenon, whereby young children simply require more information before making any speech decisions.

Most interesting, despite the ability to use prior knowledge, toddlers do not show evidence of perceptual restoration, at least in the present task. Adults typically show an increase in intelligibility when part of a signal is replaced with noise (where the missing sounds might be present but masked) over when they are replaced with silence (and are clearly missing). Children 5.5 years of age show similar results (Newman, 2004). Toddlers do not show this same pattern, at least with the stimuli and methods used here. Instead, they appear to treat words containing noise as being quite similar to those containing silent breaks. If they make a distinction at all, it appears to be in the opposite direction, with poorer identification in the noise condition than in the silence condition, as can be seen by looking across all five of the figures: In none of the figures does the average value for the noise condition exceed that for the silence condition. (Moreover, two of the four experiments reported here showed significantly better performance in the silence condition than in the noise condition.) Furthermore, this lack of the predicted difference between the noise and the silence conditions does not appear to have been due simply to ceiling effects: Although the children may have been at ceiling in some of these experiments, they clearly were not so in all of them. Indeed, all four of the experiments reported here showed relatively better performance in the clear condition than in the noise condition, even if this was not significant in all cases. (In Experiment 1, looking was 3.3 sec longer to the matching video than to the mismatching video in the clear condition, but only 0.4 sec longer in the noise condition; in Experiment 2, looking was 2.0 sec longer to the matching video than to the mismatching video in the clear condition, but only 1.5 sec longer in the noise condition; in Experiment 3, looking was 2.8 sec longer to the matching video than to the mismatching video in the clear condition, but only 1.9 sec longer in the noise condition, and in Experiment 4, looking was 2.33 and 4.23 sec longer to the matching video in the clear condition for the between and within conditions, respectively, but 2.53 and 2.86 sec longer to the matching video in the noise condition.) Since performance was not at ceiling levels, there is no reason why an increase in intelligibility in the noise condition could not have occurred. Yet even with items that adults heard as intact 100% of the time (the “lobster” and “dinosaur” tokens in Experiment 3), the children performed more poorly when noise replaced a phoneme than when silence replaced the phoneme. Thus, prior knowledge does not seem to “fill in the gaps” in toddlers’ perception; instead, toddlers appear to be more tied to the acoustic information that they receive.

It may not seem that surprising that toddlers place more emphasis on the acoustic input and less emphasis on prior

knowledge than do adults. But these results also suggest that children may not have the perceptual illusion of restoration. Although we were not testing for the illusion directly (by asking the children to distinguish explicitly whether a phoneme was present or absent), experiencing such an illusion should lead to an increase in intelligibility in the noise condition. The fact that the increase in intelligibility did not occur suggests that the illusion of restoration also was absent. Of course, it is also possible that the noise stimuli used here, although sufficient to induce restoration in adults, were simply insufficiently similar to the absent phonemes to induce restoration in young children. The latter argument would suggest that children might show the illusion of restoration in some environments but that such an illusion is simply more limited for toddlers than for adults. Perhaps restoration develops gradually, so that it begins to be present only in very limited situations. But whether it is much more limited or entirely absent, restoration clearly does not occur for toddlers in situations in which it does occur for older listeners.

There are several possible reasons for these findings. First, toddlers may not be as sensitive to prior knowledge as are adults. This would imply that there is a change in word recognition processes with development. Even when children have sufficient lexical knowledge to fill in missing information, they may not use that knowledge in real-time processing. In support of this idea, Elliott et al. (1983) have found that young children appear to be less sensitive to word frequency information than are older children and adults. In addition, older children appear to be more adept at using prior knowledge to help interpret partial acoustic information (Elliott et al., 1979).

An alternative explanation for the lack of restoration effects shown here rests on differential word learning. The children in these experiments clearly knew the target words, but they likely did not know them as well as adults do (McGregor et al., 2002) and may not have known them sufficiently well to produce restoration. That is, rather than being a sign of a reorganization in word recognition, the present results may, instead, be an indication that there is a criterion of word recognition or of lexical knowledge necessary for restoration. If so, what might this criterion be? The fact that adults show restoration for nonwords may seem to argue against a strict notion of criterion; moreover, adult listeners generally show as much (or even more) restoration for low-frequency words as for high-frequency words (Samuel, 1996a). However, Walley (1988) has argued that the differences between the size of toddlers' and adults' lexicons and the differences in their experienced word frequencies are such that prior knowledge and partial sensory information "would not constitute a very reliable basis for recognition by the young child" (p. 139). Since toddlers are far more likely to encounter novel words than are adults, it makes more sense for them to assume that items like "dina or" are simply yet-to-be-learned words, rather than misproductions of familiar items (see Cole & Perfetti, 1980, for a similar argument). In fact, there is some evidence that children

are more accepting of new information in this way. For example, children are more likely to accept a novel sentence as a grammatical form that they have not yet heard, rather than treating it as ungrammatical (Naigles, Gleitman, & Gleitman, 1993). When presented with sentences such as "The zebra brings to the lion," adults were more likely to act out the zebra bringing something to the lion, whereas 2-year-olds were as likely to interpret the sentence according to the syntactic frame (i.e., to act out the zebra going to the lion by itself). The authors suggest that the adult listeners knew the sentence was ungrammatical and changed it to become grammatical (by "adding" a thing to be brought). In contrast, young children were more willing to accept that the word could be used in a new way and, thus, extended the meaning of the verb according to its sentence frame.

A final possibility is that the difference between toddlers and adults has to do with the speed of processing. Cole and Perfetti (1980) examined children's ability to detect mispronunciations in a story and found that children as young as 4 years of age detected mispronunciations more accurately in a predictable context, demonstrating their ability to use contextual knowledge to recognize words with only partial information. However, young children (unlike adults) were not able to use that information quickly enough to influence processing on-line: Whereas adults are faster to detect mispronunciations in second syllables than in word-initial syllables, grade school children were not. Walley (1987) found that when a context was sufficiently constraining (when a picture referent was present), 5-year-old children did show such a position effect, but 4-year-old children did not. Thus, even when children have prior knowledge about lexical identity, they may not be able to access this information quickly enough to influence perception in real time.

There are thus several possible reasons for a lack of restoration in toddlers: Children may focus less on the prior knowledge that they have; their prior knowledge may be either less complete than adults' or may be accessed less quickly; or children may be more open to new information (less limited by their prior knowledge). Future research will be needed to distinguish among these different alternatives.

Regardless of the reason, the pattern of findings suggests that children's use of their prior knowledge differs from that for adults. Children recognize their first words by the time they reach 6 months of age (Tincoff & Jusczyk, 1999) and have developed a substantial pool of lexical knowledge by the age tested here. A number of studies have demonstrated that this prior knowledge can influence their perception. For example, information about statistical patterns in words influences infant perception before the end of their 1st year of life (Jusczyk, Cutler, & Redanz, 1993; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993), and 24-month-old children recognize words on the basis of partial information (Fernald et al., 2001; Swingley et al., 1999). Clearly, then, children 2 years of age are able to use prior knowledge to help their speech perception in a number of tasks. Yet they fail to show this

advantage in the manner in which adults do. Having access to information about lexical constraints does not entail using that information in all situations.

Given the present findings, what can we say about children's lexical representations? Clearly, children's representations for words are such that they can recognize words on the basis of partial information (see also Fennell & Werker, 2003; Fernald et al., 2001; Swingley et al., 1999), suggesting that they do not need all information to be present in order to identify a word. However, children also tend to perform more poorly at recognizing words that have been mispronounced than at recognizing words pronounced correctly (Swingley & Aslin, 2000, 2002); this implies that children's lexical representations contain fine phonetic detail. The present results, in which toddlers generally performed better with full word tokens than with tokens missing a phoneme, support this notion. Moreover, finding such an effect for a word such as "doggie" (which has no lexical neighbors and, thus, need not have a fully specified representation in order to be distinguished from other words) suggests that representations do not become detailed merely when there is pressure from lexical neighbors (see also Swingley, 2003). Rather, this would support the notion that lexical representations always contain phonetic detail (Fennell & Werker, 2003; Swingley, 2003; Swingley & Aslin, 2000, 2002), regardless of the other words children know.

Yet the advantage for the clear stimuli over the silent stimuli was not present in all cases. For example, whereas the children who participated in Experiment 1 showed a greater degree of appropriate looking in the clear condition than in the silence condition [$t(19) = 4.34, p < .0005$], the children in Experiments 3 and 4 showed only a nonsignificant trend in that direction. It may well be that children notice fine phonetic details only when a word is extremely well known (Fennell & Werker, 2003). According to their parents, the children in these experiments clearly knew the words "kitty" and "doggy"; the parents were less sure about their children's knowledge of the words "lobster" and "dinosaur," suggesting that the representations for these words might be less rich than those for "kitty" and "doggie." Perhaps this explains why the children's performance in the silence condition was less degraded in Experiment 3; although lexical representations may contain phonetic detail regardless of neighbors, this phonetic detail may be fully accessible only in words that are very well known (see Fennell & Werker, 2003, for more on this proposal). Alternatively, the phonetic detail may become more fully instantiated as the word is learned better; with words learned less well, some of the perceptual details may have been forgotten with time or may never have been fully learned and, thus, may not be available (Swingley, 2003; Swingley & Aslin, 2000). The present data cannot distinguish between these two proposals; nonetheless, the variability in performance in the present study may be the result of variation in how well the particular words were known by our listeners.

The children in Experiments 3 and 4 were clearly much older than those in the work by Swingley and colleagues

and by Fennell and Werker, making such comparisons chancy; however, it is likely that what matters is not the age of the child but how well known the particular word is to that child (a function both of the age of the child and of when that word was learned). It may also be the case that actual phonetic *mismatch* (as examined for younger children in Fennell & Werker, 2003; Swingley, 2003; Swingley & Aslin, 2000, 2002) is more clearly discrepant with a target word than is the mere *absence* of a phoneme (as in the present silence condition), especially for words that are known less well. If so, children may require even greater amounts of learning before being able to notice a missing phoneme than before being able to notice a discrepant one. Future research will be needed to examine this issue.

In conclusion, toddlers appear to place more reliance on acoustic information (and less on prior knowledge) than do adults or older children, even in those cases in which they have appropriate prior knowledge. They do not show perceptual restoration and need more of the signal to be intact in order to properly identify the stimulus. Future work will be needed to examine the reasons behind these developmental changes.

REFERENCES

- BASHFORD, J. A., JR., RIENER, K. R., & WARREN, R. M. (1992). Increasing the intelligibility of speech through multiple phonemic restorations. *Perception & Psychophysics*, *51*, 211-217.
- BASHFORD, J. A., JR., & WARREN, R. M. (1987). Multiple phonemic restorations follow the rules for auditory induction. *Perception & Psychophysics*, *42*, 114-121.
- BASHFORD, J. A., JR., WARREN, R. M., & BROWN, C. A. (1996). Use of speech-modulated noise adds strong "bottom-up" cues for phonemic restoration. *Perception & Psychophysics*, *58*, 342-350.
- BRAATEN, R. F., & LEARY, J. C. (1999). Temporal induction of missing birdsong segments in European starlings. *Psychological Science*, *10*, 162-166.
- COHEN, L. B., ATKINSON, D. J., & CHAPUT, H. H. (2000). Habit 2000: A new program for testing infant perception and cognition (Version 2.2.5c) [Computer software]. Austin: University of Texas, Department of Psychology.
- COLE, R. A., & PERFETTI, L. A. (1980). Listening for mispronunciations in a children's story: The use of context by children and adults. *Journal of Verbal Learning & Verbal Behavior*, *19*, 297-315.
- DALE, P. S., & FENSON, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, *28*, 125-127.
- DEWITT, L. A., & SAMUEL, A. G. (1990). The role of knowledge-based expectations in music perception: Evidence from musical restoration. *Journal of Experimental Psychology: General*, *119*, 123-144.
- ELLIOTT, L. L., BUSSE, L. A., PARTRIDGE, R., RUPERT, J., & DEGRAFF, R. (1986). Adult and child discrimination of CV syllables differing in voicing onset time. *Child Development*, *57*, 628-635.
- ELLIOTT, L. L., CLIFTON, L. A. B., & SERVI, D. G. (1983). Word frequency effects for a closed-set word identification task. *Audiology*, *22*, 229-240.
- ELLIOTT, L. L., CONNORS, S., KILLE, E., LEVIN, S., BALL, K., & KATZ, D. (1979). Children's understanding of monosyllabic nouns in quiet and in noise. *Journal of the Acoustical Society of America*, *66*, 12-21.
- ELLIOTT, L. L., HAMMER, M. A., SCHOLL, M. E., & WASOWICZ, J. M. (1989). Age differences in discrimination of simulated single-formant frequency transitions. *Perception & Psychophysics*, *46*, 181-186.
- FENNEL, C. T., & WERKER, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language & Speech*, *46*, 245-264.
- FENSON, L., DALE, P. S., REZNICK, J. S., BATES, E., THAL, D. J., & PETH-

- ICK, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, **59**, 1-185.
- FERNALD, A., SWINGLEY, D., & PINTO, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, **72**, 1003-1015.
- GOLINKOFF, R. M., HIRSH-PASEK, K., CAULEY, K. M., & GORDON, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, **14**, 23-45.
- HUNTER, M. A., & AMES, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 5, pp. 69-95). Westport, CT: Ablex.
- JUSCZYK, P. W., & ASLIN, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, **29**, 1-23.
- JUSCZYK, P. W., CUTLER, A., & REDANZ, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, **64**, 675-687.
- JUSCZYK, P. W., FRIEDERICI, A. D., WESSELS, J., SVENKERUD, V. Y., & JUSCZYK, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory & Language*, **32**, 402-420.
- KOROLEVA, I. V., KASHINA, I. A., SAKHNOVSKAYA, O. S., & SHURGAYA, G. G. (1991). Perceptual restoration of a missing phoneme: New data on speech perception in children. *Sensory Systems*, **5**, 191-199.
- MCGREGOR, K. K., FRIEDMAN, R. M., REILLY, R. M., & NEWMAN, R. M. (2002). Semantic representation and naming in young children. *Journal of Speech, Language, & Hearing Research*, **45**, 332-346.
- NAIGLES, L. G., GLEITMAN, H., & GLEITMAN, L. R. (1993). Children acquire word meaning components from syntactic evidence. In E. Dromi (Ed.), *Language and cognition: A developmental perspective* (Human Development, Vol. 5, pp. 104-140). Norwood, NJ: Ablex.
- NEWMAN, R. S. (2004). Perceptual restoration in children versus adults. *Applied Psycholinguistics*, **25**, 481-493.
- NEWMAN, R. S. (2005). The cocktail party effect in infants revisited: Listening to one's name in noise. *Developmental Psychology*, **41**, 352-362.
- ROSE, S. A., GOTTFRIED, A. W., MELLO-CARMINAR, P., & BRIDGER, W. H. (1982). Familiarity and novelty preferences in infant recognition memory: Implications for information processing. *Developmental Psychology*, **18**, 704-713.
- SAMUEL, A. G. (1981a). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General*, **110**, 474-494.
- SAMUEL, A. G. (1981b). The role of bottom-up confirmation in the phonemic restoration illusion. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 1124-1131.
- SAMUEL, A. G. (1996a). Does lexical information influence the perceptual restoration of phonemes? *Journal of Experimental Psychology: General*, **125**, 28-51.
- SAMUEL, A. G. (1996b). Phoneme restoration. *Language & Cognitive Processes*, **11**, 647-653.
- SAMUEL, A. G. (1997). Lexical activation produces potent phonemic percepts. *Cognitive Psychology*, **32**, 97-127.
- SASAKI, T. (1980). Sound restoration and temporal localization of noise in speech and music sounds. *Tohoku Psychologica Folia*, **39**, 79-88.
- SCHAFER, G., & PLUNKETT, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development*, **69**, 309-320.
- SCHULTZ-WESTRE, C. (1985). *A visual analog of phonemic restorations: Sign restoration in American Sign Language*. Unpublished doctoral dissertation, University of Wisconsin, Milwaukee.
- SWINGLEY, D. (2003). Phonetic detail in the developing lexicon. *Language & Speech*, **46**, 265-294.
- SWINGLEY, D., & ASLIN, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, **76**, 147-166.
- SWINGLEY, D., & ASLIN, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, **13**, 480-484.
- SWINGLEY, D., PINTO, J. P., & FERNALD, A. (1999). Continuous processing in word recognition at 24 months. *Cognition*, **71**, 73-108.
- TINCOFF, R., & JUSCZYK, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, **10**, 172-175.
- TROUT, J. D., & POSER, W. J. (1990). Auditory and visual influences on phonemic restoration. *Language & Speech*, **33**, 121-135.
- WALLEY, A. C. (1987). Young children's detections of word-initial and final mispronunciations in constrained and unconstrained contexts. *Cognitive Development*, **2**, 145-167.
- WALLEY, A. C. (1988). Spoken word recognition by young children and adults. *Cognitive Development*, **3**, 137-165.
- WARREN, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, **167**, 392-393.
- WARREN, R. M., & OBUSEK, C. J. (1971). Speech perception and phonemic restorations. *Perception & Psychophysics*, **9**(3-B), 358-362.
- WARREN, R. M., & SHERMAN, G. L. (1974). Phonemic restorations based on subsequent context. *Perception & Psychophysics*, **16**, 150-156.
- WARREN, R. M., & WARREN, R. P. (1970). Auditory illusions and confusions. *Scientific American*, **223**, 30-36.

NOTES

1. In our initial pilot study, we tested twenty 16-month-old toddlers. However, these toddlers did not look longer at the appropriate video for the clear condition stimuli.
2. The dog did move off the screen at one point in the original taping; this was simply edited out from the test tape, so that the final tape showed 10 min of a fairly quiescent animal.
3. Although prior work has suggested that white noise leads to more restoration than do tones for stop consonants (Samuel, 1981b), this work did not distinguish between voiced and voiceless stops. Generally, the more similar the noise to the phoneme it is replacing, the better the restoration. It is therefore unsurprising that voiced and voiceless stops would differ in this regard.
4. Parents filled out an initial questionnaire prior to the study, asking about their children's knowledge of these two words. Eighteen of the 20 toddlers were reported to know both of these words; the other 2 parents declined to answer this question.
5. That is, the toddlers were not looking around the room, as for the source of a startling noise, nor were they attempting to leave the booth or looking at their parent to suggest that they were tired of the videos.
6. Thanks to Letitia Naigles for suggesting this possibility.
7. Unfortunately, a tape containing 1 child's data was lost between the time it was initially coded and the time at which this reanalysis took place.
8. Thanks to Arty Samuel for suggesting this possibility.

(Manuscript received January 11, 2003;
revision accepted for publication July 19, 2005.)