Developmental Science

Developmental Science (2014), pp 1–10 DOI: 10.1111/desc.12182

PAPER

Why the body comes first: effects of experimenter touch on infants′ word finding

Amanda Seidl,[1] Ruth Tincoff,[2] Christopher Baker[1] and Alejandrina Cristia[3]

1. *Speech, Language and Hearing Sciences, Purdue University, USA*
2. *Department of Psychology, Bucknell University, USA*
3. *Laboratoire de Sciences Cognitives et Psycholinguistique, Centre National de la Recherche Scientifique, ENS-DEC EHESS, Paris, France*

Abstract

The lexicon of 6-month-olds is comprised of names and body part words. Unlike names, body part words do not often occur in isolation in the input. This presents a puzzle: How have infants been able to pull out these words from the continuous stream of speech at such a young age? We hypothesize that caregivers' interactions directed at and on the infant's body may be at the root of their early acquisition of body part words. An artificial language segmentation study shows that experimenter-provided synchronous tactile cues help 4-month-olds to find words in continuous speech. A follow-up study suggests that this facilitation cannot be reduced to the highly social situation in which the directed interaction occurs. Taken together, these studies suggest that direct caregiver–infant interaction, exemplified in this study by touch cues, may play a key role in infants' ability to find word boundaries, and suggests that early vocabulary items may consist of words often linked with caregiver touches.
A video abstract of this article can be viewed at http://youtu.be/NfCj5ipatyE

Introduction

Imagine a familiar parenting scene in which a mother is changing her 6-month-old's diaper. The infant kicks her feet up. They are so adorable that the mother reflexively grabs a foot and tickles it. Now imagine that this same direct caregiver–infant interaction is simultaneous with a phrase containing the word 'feet'. In this paper, we argue that such natural events may not only hold the key to the remarkable composition of 6-month-olds' lexicons, but may also challenge our views of what kinds of informational streams infants attend to when learning language.

From very early on, infants pick up on names: 'mommy', 'daddy' (Tincoff & Jusczyk, 1999), as well as their own name (Bortfeld, Morgan, Golinkoff & Rathbun, 2005; Mandel, Jusczyk & Pisoni, 1995). Their success is likely due to a number of factors, one being that these words name people who are important in the infant's social world. Names also occur frequently during parent talk, often in isolation, and may be produced with the attention-grabbing tune used when calling a person's name (vocative chant; Ladd, 1997).

Recent data show that infants' tiny vocabularies also include body part words. Indeed, 6-month-olds look longer to a video of feet than a video of a hand when hearing the word 'feet', and they display the inverse tendency upon hearing the word 'hand' (Tincoff & Jusczyk, 2012). This reliable picture-matching behavior also occurs when body parts are paired with food items (Bergelson & Swingley, 2012). This early acquisition of body part vocabulary cannot be explained in the same way as names: the vocative chant is not appropriate for them; there is no reason why feet should be spoken of more frequently in infant-directed speech than diaper, and body part words are not among the minuscule proportion of words spoken in isolation (surrounded by pauses) in infant-directed speech (Johnson, Seidl &

Address for correspondence: Amanda Seidl, Department of Speech, Language and Hearing Sciences, Purdue University, 500 Oval Dr., West Lafayette, IN 47907, USA; e-mail: aseidl@purdue.edu

Tyler, 2014). In fact, an analysis of a large corpus of speech directed to a Dutch- and German-learning 6-month-old infant (van de Weijer, 1998) reveals that while body part words comprise 11% of the nouns in the infant-directed speech (IDS) found in the corpus, they occur in isolation only 2 times. Thus, this means that out of the body part words produced, only 0.4% of these instances are in isolation, 53% of the time they are aligned to one edge, and the remaining 47% of the time they occur utterance-medially. For comparison, other nouns in the same corpus occur 10% of the time in isolation, 54% of the time aligned to an edge, and only 35% utterance-medially. If we take this corpus as representative of infants' input, then it seems that body part words are more often found sentence-medially than other nouns in IDS, and less frequently in isolation than other words.[1] Thus, the first challenge for the infant who comes to learn 'hands' and 'feet' by 6 months is to locate these words in running speech, which is difficult because there are no reliable cues to word edges (Cole & Jakimik, 1980). Even at 6 months, infants in the lab only succeed in this process of word segmentation under ideal conditions where at least one of the word's edges is given for free (e.g. it occurs at the edge of a sentence, Johnson *et al.*, 2014). This presents a puzzle: How have infants been able to pull out these words from the continuous stream of speech at such a young age?

While body part words do not have the same social and informational value as proper names, they do have a potentially crucial advantage that may license their place in the early lexicon: They refer to the infant him/herself. Since physical contact is widespread in infant–caregiver interactions (Stack & Muir, 1990), the most natural way of drawing the infant's attention towards the referent of a body part word could involve direct caregiver interaction, perhaps best exemplified by touch. While parenting style and parenting gesture use have been shown to influence language development (Masur, Flynn & Eichorst, 2005; Rowe, 2012; Rowe & Goldin-Meadow, 2009) the impact of caregiver–infant interaction exemplified by touch has yet to be explored with respect to language learning. This is a gap worthy of being filled since caregiver touch could serve to boost learning of body part words in many ways.

First, caregiver touches could serve to focus infants' attention on both the specific body part as well as other ongoing experiences. By their nature, auditory and tactile cues are less easily ignored than, for instance, visual ones, since it is easier for the young immobile infant to close her eyes or reorient her head than it is to move her body parts out of reach of the caregiver. Indeed, signers often tap their child when trying to attract their attention (Harris, Clibbens, Chasin & Tibbitts, 1989), an effective strategy that helps avoid the chance of signing in front of eyes that can be voluntarily closed. Second, caregiver touches could provide cross-modal, synchronous cues to bootstrap segmentation of the speech stream. Other work suggests that caregivers link acoustic information with action/object boundaries by e.g. moving a target object simultaneously to saying a target word (Gogate, Bahrick & Watson, 2000; Meyer, Hard, Brand, McGarvey & Baldwin, 2011). Finally, caregiver touches co-occur with close caregiver proximity, which should boost arousal. For example, we know that infants respond to their mother's proximity/smile (Goldstein, King & West, 2003) and caregiver-mediated tactile stimulation tends to be an emotionally and socially rich signal as well (Hertenstein, 2002).

Second, this research gap is worth exploring because body part words might play a key role in lexical development. Indeed, if infants are able to learn body part words early on using cues gleaned from caregiver–infant interactions, these words, by virtue of being similar in structure to other common nouns, could provide a toe-hold for the infant to gain the remainder of her vocabulary. And perhaps exploring the nature of cues provided in caregiver–infant interactions could help to explain the special status of body part words both neurologically (body part words tend not to be lost with aphasias; Kemmerer & Tranel, 2008) and cross-linguistically (they can gain the status of closed class words cross-linguistically; Heine, 1997).

Finally, previous work on cross-modal correlations in infant speech processing has focused solely on visual-auditory interactions (Hollich, Newman & Jusczyk, 2005; Jesse & Johnson, 2012; Thiessen, 2010), but in general cross-modal correlations have been found to be helpful. Furthermore, tactile information has been shown to boost speech perception in adults. Indeed, both blind and sighted adults' ability to recognize spoken syllables is enhanced by congruent tactile information (Fowler & Dekle, 1991; Gick, Jóhannsdóttir, Gibraiel & Muehlbauer, 2008; Sato, Cavé, Ménard & Brasseur, 2010). Motivated by this work, *we hypothesized that caregiver–infant interaction as exemplified by*

---

[1] Note that this corpus does appear to be representative. We also examined two corpora of English. Specifically, we examined both the Korman corpus (Korman, 1984) (a corpus directed to a 10-week-old) as well as the Soderstrom corpus (MacWhinney, 2000; Soderstrom, Blossom, Foygel & Morgan, 2008). Counts of body parts and their sentence locations in these two corpora revealed that body part words were never uttered in isolation, occurred only once sentence-initially, but were uttered sentence-medially (242 times) and sentence-finally (387 times).

caregiver touch could play a key role in early lexical development. More specifically, we tested whether caregiver touches, when reliably aligned with a spoken word form, boost 4-month-olds' memory for that word form, by using an artificial language paradigm that has been extensively used to assess the importance of word segmentation cues (Saffran, Aslin & Newport, 1996).

Experiments

In order to provide a basis for full-fledged word-knowledge by 6 months, infant word-finding skills should be boosted by simultaneous and directed caregiver–infant interactions in even younger infants. Therefore, participants in the present research were 4 months of age. To assess the value of synchronous experimenter interactive cues, infants were exposed to a continuous stream of speech with no acoustic or distributional cues to word boundaries, and only the experimenter's gestures provided information that could underline a syllable sequence. Memory for the words was assessed subsequently in a preferential listening test known as the Headturn Preference procedure (HPP; Jusczyk & Aslin, 1995). This procedure has been used reliably with infants as young as 4–6 months of age in multiple labs (Dawson & Gerken, 2009; Fernald, 1985; Hoehle, Bijeljac-Babic, Herold, Weissenborn & Nazzi, 2009; Kemler Nelson, Jusczyk, Mandel, Myers, Turk & Gerken, 1995; Johnson *et al.*, 2014; Mandel *et al.*, 1995; Nazzi, Jusczyk & Johnson, 2000; Seidl, Cristià, Bernard & Onishi, 2009) and can yield both familiarity (Thiessen & Saffran, 2003) and novelty preferences (Hay & Saffran, 2012; Saffran *et al.*, 1996) in older and younger infants. While it is not entirely clear whether infants will favor novel or familiar sequences in any task (see e.g. a discussion in Gerken, Dawson, Chatila & Tenenbaum, 2014), most researchers suggest that when infants are exposed to streams of repeated statistically recurring syllables with little complexity and much repetition, a novelty preference is more likely. Nonetheless, making a specific preference prediction is unnecessary here given that we built in an internal control, as will be explained below.

Specifically, in Experiment 1, infants were touched by an experimenter on one body part in synchrony with all instances of a particular trisyllabic syllable sequence (we call this an *Always* word). At test, we played isolated instances of this Always word, and also isolated instances of a *Nonword*, a novel arrangement of syllables that were presented in the stream, but not in the same order. A difference in listening behavior between these two types would indicate that infants can recognize the difference between familiar sequences of syllables and unfamiliar ones in this very demanding task. In fact, 7-month-olds in Curtin, Mintz and Christiansen (2005) failed to retain the statistical words present in this same stream in the absence of accompanying cues, and could only do it when acoustical cues were added.

Nonetheless, to provide an internal confirmation that a preference between the Always word and the Nonword related to a tactile boost in word finding, we also included a third trial type, which we call the *One* word. During familiarization, the experimenter consistently touched one body part for the Always word, and another body part once for every other trisyllabic sequence in the stream. Thus, two body parts were touched the same number of times, but one was touched always in co-occurrence with the same word form, and the other was touched in co-occurrence with 20 different trisyllables. Isolated tokens of one such trisyllable were also included among the test trials. Thus, the One word is used to gauge the impact of the reliability of the experimenter–infant interaction on infant behavior.

The inclusion of an Always word, a Nonword, and a One word allows us to make specific predictions. If the infant attends only to the auditory stream, she would treat the Always and the One sequences in exactly the same manner, since both had been presented the same number of times and were statistical words. She should also treat the Nonword differently from these other two test item types. Given that 7-month-olds have failed to retain the statistical words in this same stream (Curtin *et al.*, 2005), such an outcome would indicate that any caregiver interaction, even when it is unreliable, boosts statistical learning. Alternatively, if *consistent* trisyllable–body part touch synchrony helped the infant segment the exposure stream, then only the Always word should be more accessible to the infant's memory and the infant will respond differentially to this item compared to the other two types.

Experiment 1 thus investigated wholesale effects of caregiver interaction as well as reliability using direct touch. As noted above, touch is common but not the only interactive gesture caregivers make. On the contrary, much previous work has investigated cross-modal cues to word learning as instantiated in caregiver pointing to, or moving, an object. In the context of body part words, such visual referential cues would be most naturally captured through a caregiver that points or touches her own body parts. Thus, infants in Experiment 2 observed an experimenter touching her own chin or eyebrow following the exact same design as in Experiment 1. In general terms, this second group also benefits from an interactive caregiver in a highly social setting, and they also receive information that is as reliable as for infants in Experiment 1. However, these touches differ from Experiment 1's both in their modality (tactile versus visual) and the owner/perspective (self versus other), both of which could impact results. In particular, the key question is whether self-tactile and other-visual cues have the exact same role in word finding, in which case we should observe the same pattern of preferences in this follow-up.

Participants

Forty-eight typically developing full-term English-learning 4-months-olds (17 female; 12 in Experiment 2) (mean = 4.8 mos for Experiment 1; 4.5 mos for Experiment 2, min = 4.2 mos for Experiment 1 and 3.8 mos for Experiment 2, max = 5.6 mos for Experiment 1 and 5.5 mos for Experiment 2) were tested. Twelve more infants were run, but excluded due to fussiness in exposure or test (seven), experimenter error (three), or orientation times greater than 3 standard deviations off the mean (two).

Stimuli

We created an exposure string of syllables that was an adapted version of the continuous string in Curtin *et al.*, (2005). In our adaptation, all cues to word edges (stress, coarticulation, allophony) were neutralized. Thus, the only cues differentiating Always and One items in Experiments 1 and 2 were the information provided by the experimenter in a primarily tactile/visual form. To create the exposure string we recorded individual, monosyllabic units all with primary stress. Each syllable consisted of an onset consonant(s), a vowel, and a coda consonant. The identity of the coda was identical to that of the onset of the following syllable, allowing for full inter-syllable coarticulation, thus rendering coarticulation an impossible cue for segmentation. In addition, syllables with initial voiceless obstruents were recorded with a proceeding /s/ in order to avoid English allophonic aspiration cues, which may have led to the interpretation of a word boundary. This /s/ and the coda consonants were both then excised in Praat (Boersma & Weenink, 2005; a program for acoustic analysis and manipulation). For example, the portion of the exposure string '… nepokuta …' was recorded in four individual syllables 'nep', 'spok', 'skut', and 'stan', all with primary stress, and then trimmed in Praat to yield 'ne', 'po', 'ku', and 'ta' which were then spliced together. The stimuli were recorded by a young female adult speaker in a child-friendly voice. Specific tokens of syllables were chosen based on similarities in duration and pitch; amplitude was normalized for each syllable from 82 to 87 db. The syllable durations ranged from .45 to .60 s, the pitches ranged from 5.7 to 6.45 ERB (all differences were non-significant). These syllables were concatenated into a 27-syllable exposure string which played in a loop 24 times, resulting in an exposure time of 5 min and 45 s. In addition to this exposure string, we also created three trisyllabic test items. All test items were recorded in the same way as the exposure string; however, the final syllable was recorded without a coda consonant since the test items were not played in a loop, but instead were played in isolation. For example, the test item 'lepoga' was recorded as 'lep', 'spog', 'ga', with the initial /s/ and final consonants later excised in Praat.

Design

Of the three trisyllabic sequences that were used as test items, two consisted of adjacent trisyllables found in the exposure string (henceforth referred to as 'words'). Each of these test items occurred only once in the string. The exposure string, with the two test items in italics, consisted of 'gabigamunepokutanedoku*lepoga*donemutale*dobita*pomubileku' looped 24 times. We also included a Nonword test item which consisted of three syllables that occurred in the exposure string, but were not adjacent in the same order presented at test. This item was 'bipota' for all infants. Notice that all three test items were matched in terms of the frequency of their component syllables in the string. Infants were randomly assigned to one of three counterbalanced orders, determined by the co-occurrence of test word, body part touched (either knee or elbow), and frequency of touching (Table 1).

For example, in order 1, the test item 'dobita' was correlated with a consistent touching of the knee in Experiment 1 and the experimenter's chin in Experiment 2, meaning that during each repetition of the string, the knee/chin would be touched throughout the time while the sequence 'dobita' was playing. Additionally, a second body part was touched once during a control test item.[2] For example 'lepoga' in order 1 co-occurred with a single touch on the knee in Experiment 1 and a single touch on the chin in Experiment 2. In fact, this second body part was touched for every other trisyllable in the string once (for a total of 24 touches), with the only exception being the Always item. For example, in order 2 (see Figure 1), participants were

---

[2] In Experiment 1 we selected body parts which were distant, but not easy for the baby to see while in close proximity to an experimenter. In Experiment 2 we selected body parts on the face (given that infants' attention is drawn to faces) that were distant yet easily touchable.

Table 1. *Orders for Experiments 1 and 2*

Orders	dobita	lepoga	bipota
1	always elbow/eyebrow	one knee/chin	Nonword
2	one elbow/eyebrow	always knee/chin	Nonword
3	always knee/chin	one elbow/eyebrow	Nonword

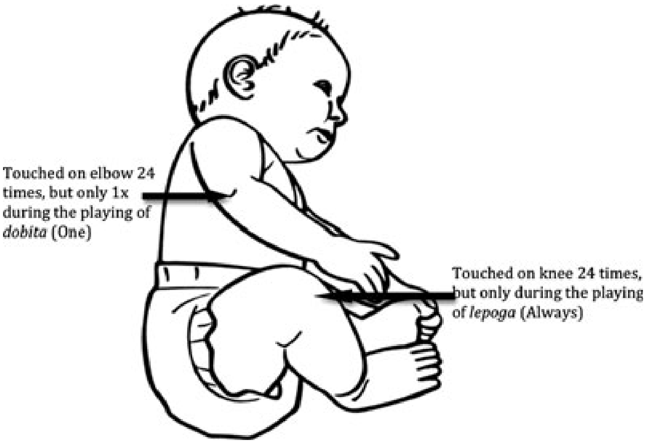

Figure 1 *Sample of infant touches in the exposure from order 2 of Experiment 1.*

touched on the elbow/eyebrow once during every trisyllabic sequence (including ‘dobita’) except for ‘lepoga’; and they were touched on the knee/chin consistently for ‘lepoga’.

Procedure

The exposure phase took place with the infant seated on her caregiver’s lap in a quiet room. The experimenter wore aviator noise-canceling headphones and sat closely facing the infant with her face at the infant’s eye-level (see Figure 2) so that the infant would not easily see the experimenter touching her knee/elbow in Experiment 1, but she could easily see the experimenter’s eyebrow/chin in Experiment 2. For Experiment 2, infants were rated on how attentive they were to the experimenter’s face on a scale of 1–5. (No infant scored lower than 3.)

A channel splitter was used so that the infant heard the exposure string described above over speakers, while the experimenter heard a sequence of tones and masking white noise on her headphones. The tones mapped directly to the commands given to the experimenter. A high tone signaled the experimenter to touch the higher body part in each experiment (elbow/eyebrow) and a low tone signaled the experimenter to touch the lower body part (knee/chin). Each contrast tone was preceded by a brief warning tone of the same height to insure accuracy and synchrony of touches. For example, in order 2, a low warning tone played 200 ms before a subsequent low tone, which lasted the duration of ‘lepoga’. This allowed the experimenter to accurately apply the touches in relation to the test items while still being blind to their identity. During the exposure process, the infant simply sat comfortably on the caregiver’s lap.[3] After finishing the exposure phase, each participant was carried across the hall to a room housing the Headturn Preference procedure (HPP; Jusczyk & Aslin, 1995).

Figure 2 *Infant and experimenter in the exposure phase of Experiment 1.*

During the HPP the infant was seated on a caregiver’s lap in a small room with lights on the front and side walls and an audiospeaker behind each side light. At test, caregivers and experimenters wore Peltor aviation headphones and listened to a combination of loud masking music and white noise so as not to influence the infant’s behavior. Each trial began with the front light flashing to attract the infant’s attention. After the infant oriented towards it, the light was extinguished and one of the two side lights began flashing. Once the infant oriented towards that light the experimenter pressed a button on a button box which directed the computer to play a sound. Looking time towards the speaker playing the sound was recorded when the infant maintained orientation within 30 degrees of the flashing light after an initial 90-degree headturn towards it. Total looking time did not include time orienting away, although during orientations away shorter than 2 consecutive seconds, the sounds and flashing lights

---

[3] An example of the procedure and both streams (to experimenter and infant) are provided as supplementary materials.

continued. When the infant oriented away from the light for more than 2 seconds the sounds and light were extinguished and the next trial began. Test sequences were presented from the single side speaker behind whichever side light was flashing on that trial and, as mentioned, both lights and sounds were contingent on the infant's head orientation. There were a total of six test trials of three types (two *Always*, two *One*, two *Nonword*) presented in two blocks with trial order and light-side fully randomized by a computer program.

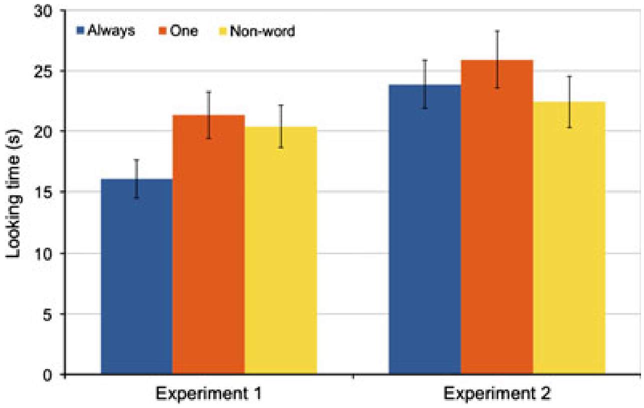

Figure 3 *Mean orientation times and standard errors in orientation times to Always, One, and Nonword test items for Experiment 1 (infant touches) and the Experimenter 2 (experimenter touches).*

Results

For Experiment 1, a repeated measures ANOVA using the factor Trial Type (Nonword vs. One Touch Word vs. Always Touch Word) showed a significant difference in looking times [$F(2, 46) = 5.105$, $p = .010$]. Infants oriented longer to the Nonwords ($M = 20.38$, $SD = 8.54$) and One words ($M = 21.31$, SD 9.42) than they did to the Always words ($M = 16.07$, $SD = 7.70$). A series of Bonferroni-adjusted paired, two-tailed t-tests using an alpha level of .0167 confirmed this pattern of differences [Nonword vs. One $t(23) = .468$, $p = .644$; Always vs. Nonword $t(23) = -3.019$, $p = .006$; Always vs. One $t(23) = -2.928$, $p = .008$]. This greater response for Nonwords was very consistent: 19/24 infants looked longer during the Nonword trials than the Always trials, which is significant via a sign test ($p < .01$).

In contrast, the same analyses in Experiment 2 failed to reveal any reliable differences in looking patterns across the three trial types [$F(2, 46) = .882$, $p = .421$; Nonword $M = 22.42$, $SD = 10.44$, One $M = 25.91$, $SD = 11.57$, Always $M = 23.88$, $SD = 9.80$; Bonferroni-adjusted t-tests, all $p > .200$]. Only 8 of the 24 infants looked longer during Nonword than Always trials ($p > .1$).

As is evident in Figure 3, results appear different across these two experiments. We focused on the differences between looking times to Nonwords and Always words because both transitional probabilities and the consistent interactive cues were aligned in predicting differential processing for Nonword and Always words, and because previous work using the same design has employed this contrast (Curtin *et al.*, 2005). Examining the Always–Nonword looking time difference scores showed that infants in Experiment 1 ($M = -4.31$, $SD = 7.00$) responded significantly differently from infants in Experiment 2 ($M = 1.46$, $SD = 11.87$) based on an independent samples t-test [$t(46) = -2.05$, $p = .046$]. This result was confirmed by a Mann Whitney U test [$U(46) = 184$, $z = -2.14$, $p = .032$] and a chi-squared test on the number of subjects [Experiment 1 19/24, Experiment 2 8/24, $\chi^2(1, N = 48) = 10.24$, $p = .01$]. In contrast, comparing the Always–One looking time difference scores was non-significant in the t-test [Experiment 1 $M = -5.37$, $SD = 8.78$, Experiment 2 $M = -2.03$, $SD = 11.35$; $t(46) = -1.14$, p = .26] and Mann Whitney U test [$U(46) = 223$, $z = -1.34$, $p = .18$] though there was a significant effect when comparing the number of subjects [Experiment 1 18/24, Experiment 2 11/24, $\chi^2(1, N = 48) = 4.27$, $p = .039$].[4]

Notice that it is unlikely that the diversity in results was due only to tactile stimulation being more arousing, at least in the test phases of the experiments. First, looking times are overall higher in Experiment 2, consistent with the opposite interpretation. Second, within the subset of infants who were highly attentive to the experimenter's face (i.e. scored a 4 or 5 on attentiveness to the experimenter's face as described in the Methods) there was still no difference in looking times between the Always, One, and Nonwords.

Together, these results suggest that there is clearly a boost in word finding in the presence of self-/infant-directed tactile cues (Experiment 1); and that this boost cannot be completely reduced to the factors that are in common between both experiments, namely the presence of synchronous cross-modal information and a highly social setting with an experimenter producing interactive

---

[4] Two separate linear mixed models were declared with looking time in each trial as the dependent measure; Experiment, Trial Type, and their interaction as fixed factors; and random intercepts for infant and trial number in one model, and infant, trial number, and gender in the other. Crucially, the inclusion of a random intercept for gender did not change the estimates for Trial Type or the interaction term at all (only the estimate for Experiment changed slightly).

gestures. We return to this interpretation in the discussion.

Discussion

For a word to earn a place in the early vocabulary, infants must be able to find that word form in fluent speech, have some knowledge of a referent on which to map the word form, and track the co-occurrences of word form and referent in order to build this connection. Experiment 1 provided a lab-based demonstration that experimenter-provided touches may facilitate infants' abilities to find words in continuous speech. When deprived of all acoustic and distributional cues, young infants were able to break up the acoustic stream when provided with frequent, synchronous and consistent, highly interactive touch cues. The key conclusion from Experiment 1 is simply that these experimenter-provided tactile cues helped the infants to perform a very difficult word-finding task, a conclusion that was unexpected prior to the evidence provided here. The precise mechanisms underlying these effects deserve much further exploration.

We took one small, first step in the direction of exploring the mechanisms behind these effects in Experiment 2, which is, in structure and general description, very similar to Experiment 1. In this follow-up study, however, word forms were paired with frequent and consistent visual/other-directed cues in the same social situation, thus incorporating both the informational and social value of the experience without the use of self-directed touch. However, at test, infants in Experiment 2 showed no evidence of recognition of the words from the exposure phase of the experiment. Thus, it appears that co-occuring experimenter-provided tactile stimulation to the infant herself is especially suited to bootstrap 4-month-olds' word finding, and that synchronous multi-modal body-related information in a rich social setting is not *sufficient* to promote word finding in this specific setting.

Aside from this possibility, then, what are the precise mechanisms that enabled infants to locate the frequently occurring trisyllables in Experiment 1? We know very little about the cognitive mechanisms triggered by touch across the lifespan and even less about how these mechanisms may interact with language (for a review of touch across the lifespan see e.g. Gallace & Spence, 2010). We have already discarded the possibility that there is a direct increased arousal effect on learning, since there is evidence to the contrary in looking times (longer in Experiment 2 than Experiment 1). Naturally, this option could be revisited using psychophysiological methods that more directly measure arousal. In addition, there remain several open explanations based on general effects of self-directed touch. To begin with, it is possible that self-directed touches provide a richer informational stream than touches to other individuals.[5] This predicts that the 'same' touch signal is encoded more faithfully when it is self-directed, a prediction that perhaps can only be assessed using neuroimaging. Moreover, self-directed touch may lead to better participant behavior, since adults who are touched are more compliant (Guéguen & Fischer-Lokou, 2003). This latter option could be tested behaviorally, by describing infants' facial expressions and movement during touch and non-touch bouts. Other hypotheses relate infants' differential performance in the two experiments to their readiness to learn from specific signals in the context of a language task. Indeed, it may be that infants at this age pay special attention to self-directed, consistent touch *because* they are already learning body parts in their every day life, which renders them particularly sensitive to caregivers' touches to their own body parts. It follows that younger and older infants may not behave in the same way, and that they may not similarly extract and remember non-linguistic sequences that co-occur with the touch. In sum, to tease apart these many hypotheses, future studies should include behavioral, psychophysiological, and neuroimaging measures, gathered in paradigms focusing on language, on memory, and on social interactions. We expect that a multi-disciplinary, extensive research approach will be necessary to shed light on the pathways subtending the improved performance.

While we did not predict any particular direction of preference since both novelty and familiarity have been found in similar paradigms (Hay & Saffran, 2012; Saffran *et al.*, 1996; Thiessen & Saffran, 2003) and with infants of similar ages (Johnson *et al.*, 2014), in the test phase of Experiment 1 infants showed what can be viewed as a novelty preference for the *Nonwords* and the *One* words relative to the *Always* words. In other words, after 5 minutes of exposure to a speech stream coupled with reliable touch, infants had gained enough sensitivity to the Always words that they showed a lack of preference for them at test. Such novelty effects are often encountered when tasks are overly simple (Hunter & Ames, 1988); for instance, in word segmentation, this has been reported for target age groups who have mastered word segmentation in a particular environment (Seidl & Johnson, 2006, 2008). Taking this novelty effect to heart, it appears that the addition of experimenter-provided tactile cues greatly facilitated finding

---

[5] We thank an anonymous reviewer for pointing this out.

these ‘words’. This is not to say that only tactile cues could yield such effects, as they clearly have been obtained using monomodal cues (to give just one example, that is the preference obtained in Curtin *et al.*, 2005, which uses the exact same string design but bootstraps infant word finding through lexical stress). Notice in addition that the looking times to the Nonword and One word were similar in our Experiment 1. Although the transitional probabilities within the One word (0.33) were higher than those in the Nonword (0), the string was set up such that there are 25 consistent trisyllables, all with 0.33 transitional probabilities. In such a setting, and in the absence of useful additional cues, even 7-month-olds fail to retain trisyllables and show no preference between these items and the Nonwords (Curtin *et al.*, 2005). Thus, it is possible that such young infants require stronger transitional probability cues and/or disambiguating cues (such as touch in Experiment 1 and stress in Curtin *et al.*, 2005).

While touch could promote word finding, it remains to be shown that this information is present in the infant’s natural environment. We are currently exploring how caregivers’ use of acoustic and tactile cross-modal cues differs when teaching body part versus object words. Preliminary results suggest that caregivers underline *object* words acoustically, by producing them with a longer duration and larger pitch ranges. In contrast, tactile cues tend to favor body part words, in terms of quantity, consistency of location, and overall duration of concomitant touch. Nearly half of the naming actions for body parts were accompanied by a touch, and about 75% of these touches were to the same broad location on the infant’s body. Temporal alignment was not an overly consistent cue, as the onset/offset of naming and touching actions were most often *not* temporally aligned for either object or body part words. These preliminary results fit well with those in Experiment 2 by suggesting that the advantage does not come from strict informational synchrony. Thus, we hypothesize that the principal locus of the body part advantage in real life lexical acquisition is actually the quantity and quality of touches concomitant to naming (quantity has also been found to be relevant in Bergelson & Swingley, 2013).

Our focus in this study was to investigate realistic mechanisms through which young infants may begin to incorporate body part words into their vocabulary, and thus future work should shed light on what specifically has been segmented and how these segmented words may relate to mapping sound sequences to meanings. Safely anchored in early, multimodal experience, body part words are so central to both individuals and communities that they cannot be easily erased in aphasias (Kemmerer & Tranel, 2008) and they have a place in ‘core vocabulary’ used to measure historical language change (Campbell, 2004). These words could also play a key role in the earliest stages of infant language development. Unlike the other early lexical items (proper names), body part words are common nouns, and could provide infants with crucial information as to how most nouns behave (Tincoff & Jusczyk, 2012). The present work has documented several facets of a set of mechanisms that could help 4-month-olds to these lexical stepping stones. Indeed, very young infants employ a heretofore largely ignored sense when processing speech, which their caregivers may intuitively exploit when talking about the names of body parts to their infants during close interactions.

Acknowledgements

This work was partially supported by funds from Purdue University and NSF 0843959 to AS, and ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL* to AC. Help was provided by Caroline Wade and Yuanyuan Wang. We thank the many caregivers and infants who participated. We also thank Elika Bergelson, Elizabeth Johnson, David Kemmerer, Keith Kluender, and Sophia Stone for comments on an earlier draft of this work as well as Michael Tyler and Sho Tsuji for help with the corpus data.

References

Bergelson, E., & Swingley, D. (2012). At 6 to 9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences, USA*, **109**, 3253–3258.

Bergelson, E., & Swingley, D. (2013). The acquisition of abstract words by young infants. *Cognition*, **127**, 391–397.

Boersma, P., & Weenink, D. (2005). Praat: Doing phonetics by computer (version 5.3.45) [Computer Program]. (retrieved 15 April 2013 from http://www.praat.org/)

Bortfeld, H., Morgan, J., Golinkoff, R., & Rathbun, K. (2005). Mommy and me: familiar names help launch babies into speech stream segmentation. *Psychological Science*, **16**, 298–304.

Campbell, L. (2004). *Historical linguistics: An introduction*. Edinburgh: The MIT Press.

Cole, R.A., & Jakimik, J. (1980). A model of speech perception. In R. Cole (Ed.), *Perception and production of fluent speech* (pp. 133–163). Hillsdale, NJ: Erlbaum.

Curtin, S., Mintz, T.H., & Christiansen, M.H. (2005). Stress changes the representational landscape: evidence from word segmentation. *Cognition*, **96**(3), 233–262.

Dawson, C., & Gerken, L. (2009). From domain-generality to domain-sensitivity: 4-month-olds learn an abstract repetition rule in music that 7-month-olds do not. *Cognition*, **111**, 378–382.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, **8**, 181–195.

Fowler, C.A., & Dekle, D.J. (1991). Listening with eye and hand: cross-modal contributions to speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, **17**, 816–828.

Gallace, A., & Spence, C. (2010). The science of interpersonal touch: an overview. *Neuroscience and Biobehavioral Reviews*, **34**, 246–259.

Gerken, L., Dawson, C., Chatila, R., & Tenenbaum, J. (2014). Surprise! Infants consider possible bases of generalization for a single input example. *Developmental Science* DOI: 10.1111/desc.12183.

Gick, B., Jóhannsdóttir, K., Gibraiel, D., & Muehlbauer, J. (2008). Tactile enhancement of auditory and visual speech perception in untrained perceivers. *Journal of the Acoustical Society of America*, **123**, EL72–EL76.

Gogate, L., Bahrick, L., & Watson, J. (2000). A study of multimodal motherese: the role of temporal synchrony between verbal labels and gestures. *Child Development*, **71**, 878–894.

Goldstein, M.H., King, A.P., & West, M.J. (2003). Social interaction shapes babbling: testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences, USA*, **100**, 8030–8035.

Guéguen, N., & Fischer-Lokou, J. (2003). Another evaluation of touch and helping behaviour. *Psychological Reports*, **92**, 62–64.

Harris, M., Clibbens, J., Chasin, J., & Tibbitts, R. (1989). The social context of early sign language development. *First Language*, **9**, 81–97.

Hay, J.F., & Saffran, J.R. (2012). Rhythmic grouping biases constrain infant statistical learning. *Infancy*, **17**, 610–641.

Heine, B. (1997). *Cognitive foundations of grammar*. Oxford: Oxford University Press.

Hertenstein, M.J. (2002). Touch: its communicative functions in infancy. *Human Development*, **45**, 70–94.

Hoehle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). Language specific prosodic preferences during the first year of life: evidence from German and French infants. *Infant Behavior and Development*, **32**, 262–274.

Hollich, G., Newman, R., & Jusczyk, P.W. (2005). Infants' use of synchronized visual information to separate streams of speech. *Child Development*, **76**, 598–613.

Hunter, M.A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research*, **5**, 69–95.

Jesse, A., & Johnson, E.K. (2012). Prosodic temporal alignment of co-speech gestures to speech facilitates referent resolution. *Journal of Experimental Psychology: Human Perception and Performance*, **38**, 1567–1581.

Johnson, E.K., Seidl, A., & Tyler, M. (2014). The edge factor in early word segmentation: utterance-level prosody enables word form extraction by 6-month-olds. *PLoS ONE*, **9**(1), e83546.

Jusczyk, P.W., & Aslin, R. (1995). Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology*, **29**, 1–23.

Kemler Nelson, D., Jusczyk, P., Mandel, D., Myers, J., Turk, A., & Gerken, L. (1995). The head-turn preference procedure for testing auditory perception. *Infant Behavior and Development*, **18**, 111–116.

Kemmerer, D., & Tranel, D. (2008). Searching for the elusive neural substrates of body part terms: a neuropsychological study. *Cognitive Neuropsychology*, **25**, 601–629.

Korman, M. (1984). Adaptive aspects of maternal vocalizations in differing contexts at ten weeks. *First Language*, **15**, 44–45.

Ladd, R. (1997). *Intonational phonology*. London: Cambridge University Press.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mandel, D., Jusczyk, P., & Pisoni, D. (1995). Infants' recognition of the sound patterns of their own names. *Psychological Science*, **6**, 314–317.

Masur, E.F., Flynn, V., & Eichorst, D. (2005). Maternal responsive and directive behaviours and utterances as predictors of children's lexical development. *Journal of Child Language*, **32**, 63–91.

Meyer, M., Hard, B., Brand, R.J., McGarvey, M., & Baldwin, D.A. (2011). Acoustic packaging: maternal speech and action synchrony. *IEEE Transactions on Autonomous Mental Development*, **3**, 1–9.

Nazzi, T., Jusczyk, P., & Johnson, E. (2000). Language discrimination by English-learning 5-month-olds: effects of rhythm and familiarity. *Journal of Memory and Language*, **43**, 1–19.

Rowe, M.L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, **83**, 1762–1774.

Rowe, M.L., & Goldin-Meadow, S. (2009). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science*, **323**, 951–953.

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science*, **274**, 1926–1928.

Sato, M., Cavé, C., Ménard, L., & Brasseur, A. (2010). Auditory-tactile speech perception in congenitally blind and sighted adults. *Neuropsychologia*, **48**, 3683–3686.

Seidl, A., Cristià, A., Bernard, A., & Onishi, K.H. (2009). Allophones and phonemes in infants' learning of sound patterns. *Language Learning and Development*, **5**, 191–202.

Seidl, A., & Johnson, E. (2006). Infant word segmentation revisited: edge alignment facilitates target extraction. *Developmental Science*, **9**(6), 565–573.

Seidl, A., & Johnson, E. (2008). Perceptual factors influence infants' extraction of onsetless words from continuous speech. *Journal of Child Language*, **35**, 1–24.

Soderstrom, M., Blossom, M., Foygel, R., & Morgan, J. (2008). Acoustical cues and grammatical units in speech to two preverbal infants. *Journal of Child Language*, **35**, 869–902.

Stack, D.M., & Muir, D.W. (1990). Tactile stimulation as a component of social interchange: new interpretations for the still-face effect. *British Journal of Developmental Psychology*, **8**, 131–145.

Thiessen, E. (2010). Effects of visual information on adults' and infants' auditory statistical learning. *Cognitive Science*, **34**, 1092–1106.

Thiessen, E., & Saffran, J. (2003). When cues collide: use of statistical and stress cues to word boundaries by 7- and 9-month-old infants. *Developmental Psychology*, **39**, 706–716.

Tincoff, R., & Jusczyk, P.W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, **10**, 172–175.

Tincoff, R., & Jusczyk, P.W. (2012). Six-month-olds comprehend words for parts of the body. *Infancy*, **17**, 432–444.

van de Weijer, J. (1998). *Language input for word discovery.* Nijmegen: Max Planck Series in Psycholinguistics 9.

Received: 10 August 2013
Accepted: 29 January 2014