1. Introduction

Children’s language acquisition appears effortless, and advances enormously between the first word period, at about 12 months, to the stage at which complete and often complex sentences are produced, at about 3 years. However, this apparently miraculous process may in fact rely on knowledge that the child has begun to accumulate even before birth. Indeed, work on infant speech perception in the last 40 years suggests that infants begin to attune to their native language during the first year of life. In this early attunement, it has been proposed that prosody plays a fundamental role, directly or indirectly advancing phonological, lexical, and morphosyntactic knowledge (Morgan and Demuth, 1996), as represented on Figure 1 and explained below. It should be noted that we focus on the ways that prosody could impact infants’ long-term representations, but obviously, as infants gain this knowledge, both prosody and emergent categories will affect online speech processing, as argued elsewhere (Christophe et al., 2008).

Like language, speech does not have a flat structure, but instead is organized into hierarchical prosodic units, and experimental work shows that this aids young infants segmentation of the speech stream (circle 1 on Figure 1). For instance, they prefer that the spoken signal be broken up at prosodic junctures than between junctures (Kemler Nelson et al., 1989; Soderstrom et al., 2003), and they encode prosodic units but not non-units (Nazzi et al., 2000b; Seidl, 2007; Seidl and Cristià, 2008). In addition, since prosodic boundaries are akin to perceptual boundaries, infants’ recognition of wordforms is constrained and improved through alignment with prosodic boundaries (circle 2; Gout et al., 2004; Seidl and Johnson, 2006). Furthermore, computational modeling suggests that relying
Figure 1: Direct and indirect pathways in which prosodic knowledge may advance language development. Representations are encoded in bold face; smaller font size indicates that a representational level is likely subsidiary of others; e.g., both phonotactics and rhythm are probably part of phonological knowledge, and word classes may be integrated in the lexicon. Circles 1-7 indicate areas on which there exists experimental (human or modeling) evidence, as detailed in the main text. Dashed lines represent possible long-term continuity between the rudimentary forms of the lexicon and morphosyntax hypothesized to exist in infancy, and the fuller fledged language in childhood. The present study investigates a longitudinal connection between a measure of sensitivity to prosodic structure in online speech processing (box marked with 1), and lexical and morphosyntactic development in early childhood (bottom square).

on these boundaries reduces the load involved in learning phonotactics, which in their stead further facilitate wordform recognition (circle 3; Brent and Cartwright, 1996). As a result of these processes operating in speech processing over time, attention to prosody should impact the formation of a proto-lexicon, a repertory of wordforms. Additionally, prosodic units delimit the domain within which word order is encoded (circle 5; Mandel et al., 1996), a limitation that should impact the reconstruction of syntactic skeletons, as detailed below.

A second line of research has shown that prosody may also provide infants with perceptual cues to word categories (circle 4 on Figure 1). Indeed, corpora analyses show that prosodic cues differ between lexical and functional categories (e.g., Shi et al., 1998) and even among different lexical categories (such as nouns versus verbs; e.g., Kelly, 1996). Shortly after birth, infants can use these perceptual cues to discriminate categorically between content and function words (Shi et al., 1999), and by 6 months these cues guide infants’ attention towards content words (Shi and Werker, 2001). It is conceivable that, by clustering wordforms in
the proto-lexicon, infants could develop rudimentary word classes; or that they can use universal perceptual cues to recognize word classes that are innately specified. In either case, tagging wordforms in terms of word classes could also contribute to proto-syntactic analyses, as follows.

Finally, prosody could contribute to rudimentary morphosyntactic descriptions in several indirect ways. To begin with, infants tune to their language’s native rhythm early on (circle 6 on Figure 1; e.g., Nazzi et al., 2000a) and typological research suggests that there are rhythmic cues associated with basic syntactic parameters, such as word order (circle 7; Christophe et al., 1998; Nespor et al., 1996). Additionally, if wordforms are tagged in terms of word type, this could also provide infants with information regarding their native language’s syntax. For instance, if function words often occur at the left edge of intonational phrases, the language is likely head-initial; if, in addition, verbs often precede nouns within the same phonological phrase, the language could have subject-verb-object typical word order.

In summary, considerable research documents nearly each and every step of the plausible processes by which prosody could contribute to the segmentation of words, their storage and classification, and to early morphosyntactic analyses. In view of the central role that prosody could play in natural language acquisition, we predicted that individual variation in sensitivity to prosody in early infancy would be reflected in language outcomes in toddlerhood. This prediction was based on a wealth of research documenting the predictive value of infant perceptual measures (Colombo et al., 2008). More specifically, previous work has found that language outcomes could be predicted from each one of the following measures gathered in the first year of life: auditory acuity (Benasich and Tallal, 2002; Molfese and Molfese, 1985), cognitive skills (Rose et al., 2009), brain responses to frequent lexical stress patterns (Weber et al., 2005), discrimination of native and non-native sound categories (Kuhl et al., 2008; Rivera-Gaxiola et al., 2005), and success in word segmentation tasks (Newman et al., 2006). In view if these results, we attempted to control for at least some of the likely confounding variables by also gathering a measure of cognitive abilities from a substantial portion of the infants. In addition, one previous study has shown that measures of recognition of prosodic units from fluent speech do not predict later language (Newman et al., 2006). Therefore, we measured sensitivity to prosody through a simple preference task, which may be less cognitively demanding, thus allowing us to bypass any individual variability involved in recognition processes that could obscure the effect of sensitivity to prosody.

2. Experiment

This experiment tested the hypothesis that individual variation in a measure of prosodic sensitivity gathered in infancy would predict language outcomes in toddlerhood. Three measures were gathered at 6 months of age. The index of
sensitivity to prosody was derived from two attentional measures in response to 2 versions of the same sequence of words. Both sequences were extracted from fluent speech, but only one of them had been uttered as a well-formed prosodic unit. These two types of stimuli have frequently been used in studies focusing on recognition of prosodic units (Nazzi et al., 2000b; Seidl, 2007; Seidl and Cristià, 2008).

Another measure was gathered with the aim of capturing individual variation related to general cognitive abilities. To this end, the Novelty score that results from a task known as Visual Recognition Memory (VRM) was chosen for three reasons. First, it has been studied extensively and the measure of individual differences it yields is reliable and relatively stable (Rose and Feldman, 1990). Second, its predictive value for both cognition and language has been repeatedly demonstrated (e.g., Bornstein and Sigman, 1986; Rose et al., 2009). Finally, success in the task does not depend on a single cognitive skill, but on several basic building blocks of cognition (Rose et al., 2003) and possibly also basic perceptual abilities (since the measure correlates with auditory acuity; e.g., Benasich and Tallal, 1996). Thus, we hoped this task would provide a control for variation related to the cognitive skills involved in performing any laboratory-based task (including those involved in our prosody measure).

The outcome measure was the MacArthur-Bates Communicative Developmental Inventory: Words and Sentences (CDI, Fenson et al., 1993), frequently used in studies investigating the predictive value of infant perception measures (e.g., Newman et al., 2006; Rivera-Gaxiola et al., 2005). The model shown on Figure 1 predicts prosodic processing to relate to vocabulary size on the basis of its contribution to word recognition; to lexical development, due to its role in word categorization; and to grammatical development, through the joint effect of prosodic cues to syntax, the delimitation of syntactically relevant domains, and the specification of word types. These predictions general were instantiated in the specific hypotheses stated in the Outcome section below.

2.1. Measures

**Novelty score.** The infant was seated on a caregiver’s lap in front of a large screen, onto which images were projected; in short, this measure is a ratio of preference for novel images in a pairwise comparison. More specifically, following Rose et al. (2009), infants were shown up to 9 ‘problems’, each consisting of a familiarization to an image, and a test where the familiar and a novel images are shown at the same time. Each trial began with an attention-getter (a green square with a black spot that appeared and disappeared at regular intervals, and non-speech sound tracks which varied across problems). When the infant fixated on the screen, visual stimuli were projected to the left and right of the infant’s visual field until she looked away for more than 2 seconds. During familiarization, the images to the right and left were the same (the same black-and-white photograph of a face in 5 of the problems, or a colorful set of geometric stimuli in the
other). Familiarization ended when the infant accumulated a fixed exposure time by looking at either of the sides (20 s for the faces, 10 s for the geometric stimuli). During test, the familiarization image continued to be projected on one of the sides, whereas a new stimulus appeared on the other side. There were 2 test trials in each problem, with side of the new image counterbalanced across them. After the experiment was completed, looking times during test were coded offline, from a digitized version at 30 frames per second, and a novelty score was calculated as the looking time to the novel image divided by the total looking time to either one. It should be noted that in order to avoid affecting infants’ performance in the prosodic processing task, the VRM portion of the test was stopped if the infant began to seem tired, such that not all infants went through all 9 problems.

Prosodic sensitivity. The prosodic sensitivity index resulted from a ratio of two measures gathered using a variant of the Head-turn Preference procedure (Jusczyk and Aslin, 1995). In this procedure, the infant sits on his/her caregiver’s lap in a three-sided booth, which has a green light at the front of the infant and red lights at each side. An experimenter observes the infant through a peep-hole and codes her headturns. Trials start with the green light blinking. When the infant faces forward, the green light is extinguished, and one of the side lights begins to blink. The auditory stimulus is played from the moment the infant orients to this side light until she orients away for more than two seconds. The time spent oriented towards the light is considered an index of attention to the auditory stimuli played. These stimuli consisted in 2 versions of the same sequence of words, played in 2 separate trials, whose order was randomized. One of the versions was a well-formed prosodic unit, as it had been uttered as an intonational phrase. The other version had been uttered as part of two different intonational phrases. Given that the latter spanned a prosodic boundary, it did not have the prosodic envelope corresponding to a natural prosodic unit, with this being the only characteristic that made the two stimuli different. Therefore, our measure of sensitivity to prosody was calculated on the basis of the ratio of the looking time to the well-formed unit divided by the looking time to the non-unit; the logarithm was applied in order to normalize the distribution of this measure. A positive ratio indicates longer looking time to the well-formed version, the expected pattern of preference.

Language outcome. When infants were about 24 months of age, parents were invited to complete a CDI. This checklist yields a measure of total vocabulary size, which we predicted to be larger in children who had had a positive prosodic ratio. In order to assess the possible relationship with lexical development, we also calculated number of words produced in specific syntactic categories, as follows. Items in the sections Animals, Vehicles, Toys, Food, Clothing, Body parts, Household, Furniture, Outside, Places, and People were considered nouns; those in Action and Descriptive were considered predicates; those in Time, Pronouns, Question, Preposition, Articles, Helping, and Connecting were summed into function words; the remaining two categories were hard to classify syntactically (Sounds, including words like ‘moo’, and Games), were placed in the Other
Table 1: Mean (min, max) within each group in terms of age at infant test (months); birth weight (lb); maternal years of education; outcome in the infant measures, and age at CDI completion (months).

<table>
<thead>
<tr>
<th>Group</th>
<th>Age Infant</th>
<th>Weight</th>
<th>Education</th>
<th>Novelty score</th>
<th>Prosody ratio</th>
<th>Age CDI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>6.06 (5.49, 6.64)</td>
<td>7.6 (6.06, 8.75)</td>
<td>17 (13, 22)</td>
<td>.571 (.405, .656)</td>
<td>-.248 (-.658, -.005)</td>
<td>24.23 (23.67, 25.48)</td>
</tr>
<tr>
<td>Positive</td>
<td>6.18 (5.46, 6.64)</td>
<td>7.06 (5.69, 8.69)</td>
<td>15.91 (12, 20)</td>
<td>.628 (.457, .791)</td>
<td>.317 (.002, .766)</td>
<td>24.15 (23.74, 24.66)</td>
</tr>
</tbody>
</table>

category. Our specific hypothesis was that infants with more advanced prosodic abilities (positive ratio) would be ahead of the other children particularly in ‘hard’ words, such as predicates and function words. Finally, indices of grammatical development considered were use of inflections, irregulars, overregularizations, and Mean Length of Utterance (MLU). Responses to section IIA (use of plural, possessive ’s, -ing, and -ed) were given a score of 0 if the child never used them; 1 if sometimes; 2 if often. The sum of these scores was considered as an index for use of inflections. Use of irregular nouns and verbs, and overregularization in the same two categories were also summed into separate scores. Finally, a measure of MLU was calculated by counting the number of morphemes in the longest sentence examples provided by caregivers, and averaging them. We expected children in the positive group to have higher scores on all of these grammatical measures.

2.2. Participants

A total of 44 infants were tested on the VRM task followed by the prosody task, and an additional 47 completed on the prosody task only; all were given a book or toy for their participation. Data from 12 were excluded for the following reasons: Failure to complete the tasks (1); preterm delivery (4); more than one language exposure at home (6); and data loss (1). Caregivers of 25 infants responded to the request for CDIs at 24 months; 1 child was excluded for having an extremely small vocabulary (3 words). While it is unfortunate that outcome data was only available from 30% of the infants that had been tested, a sample size of 24 is comparable to published longitudinal studies on speech predictors of language: 13 in Tsao et al. (2004); 17 in Liu et al. (2009); 24 in Hurtado et al. (2008); between 22 and 47 in Newman et al. (2006).

Children were classified on the basis of their infant prosody ratio into 2 groups: Positive (5 boys and 6 girls; novelty scores were available for 9 of the infants) and Negative (6 boys and 7 girls; novelty scores were available for 9 of the infants). There were neither significant nor marginal differences across the two groups in terms of age of infant testing, birth weight, maternal education, novelty scores, or age at which the caregiver filled out the CDI (all ps>.15; means and ranges are shown on Table 2.2). Additionally, the two groups did not differ in the number of VRM problems completed (positive: 5.4, negative: 6.2; t(17)=.6).
Table 2: Mean (SE) of each group of children in terms of vocabulary size, frequency of production of inflected, irregular, overregularized forms, and MLU.

<table>
<thead>
<tr>
<th>Group</th>
<th>Vocabulary</th>
<th>Inflections</th>
<th>Irregulars</th>
<th>Overregularized</th>
<th>MLU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>192 (114)</td>
<td>1.58 (1.24)</td>
<td>1.92 (2.15)</td>
<td>0.75 (2.01)</td>
<td>3.7 (1.73)</td>
</tr>
<tr>
<td>Positive</td>
<td>328 (166)</td>
<td>3.2 (1.67)</td>
<td>3.8 (3.26)</td>
<td>3 (3.83)</td>
<td>4.07 (1.29)</td>
</tr>
</tbody>
</table>

2.3. Results

Children with a positive ratio had larger vocabulary sizes \( t(23)=2.26, p<.04 \); all tests are two-tailed, unequal variances. This pattern was not particularly larger for ‘hard’ words [Predicates: \( t(23)=2.05, p=.05 \); Function words: \( t(23)=2.22, p<.04 \)] than for Nouns [\( t(23)=2.23, p<.04 \)]. The difference across the groups did not reach significance for the Others category [\( t(23)=1.67, p=.11 \)]. As shown in Figure 2, these results hold for both boys and girls. With respect to the grammatical subsections, children with positive ratios produced inflected words significantly more often \( t(21)=2.46, p<.03 \); 1 caregiver in each group did not respond. Although the average score was higher for all the other grammatical categories, these differences did not approach significance (all \( p>.11 \); 1 caregiver in each group did not respond to any of them, and 1 additional caregiver in each group did not provide examples for the MLU calculation). Averages and standard deviations for all outcome measures are given on Table 2.3.

![Figure 2](image)

Figure 2: Total vocabulary size and number of words in each syntactic category reported by parents of boys (grey) and girls (white) who had negative (plain) or positive (dotted) ratios in infancy. Error bars represent standard error.
3. Discussion

Results supported the prediction that variation in prosodic processing in early infancy predicts later language acquisition. While these results are compatible with an interpretation whereby prosodic processing helps language advancement, there are other explanations that ought to be considered.

First, any measure is likely to tap not only the construct under study, but also attention and general cognitive skills. In order to guard against the possibility that variability in general attention or arousal drove any effects found, we used a ratio of attention to 2 comparable speech samples, differing only in terms of their prosody. Additionally, we collected a third measure, VRM, which is a stable and reliable infant predictor of childhood IQ (even better than Bayley’s Developmental Scales; see e.g., Rose et al., 2003). The two groups of infants did not differ significantly along this dimension. Therefore, it is unlikely that the relationship found between attention to clausal prosody and lexical development is mediated by general attention or processing skills.

A second possibility is that the link between prosodic processing and language acquisition is due to a lurking variable that affects both predictor and outcome, such as characteristics of the child’s linguistic environment. Indeed, previous research documents complex relationships between children’s environment and their vocabulary development (e.g., Hoff, 2003; Tamis-LeMonda et al., 2004). For example, it could be the case that (1) some caregivers exaggerate prosody in speech to their infant; (2) exaggerated prosody in caregivers’ speech positively affects infants’ prosodic processing; (3) the same caregivers tend to produce more variable word types when their child is older. While this is a fascinating possibility, there is unfortunately no research documenting any of these 3 steps. Therefore, it would be of great interest to investigate the relationship between lexical and acoustical characteristics of caregivers’ speech, their long-term stability, and their joint and independent impact on child speech perception and language development.

In the meantime, we can entertain the possibility that prosody does affect language development in the long term. As mentioned in the Introduction, there is a considerable amount of work supporting the impact of prosody on language advancement through computational models and controlled laboratory-based experiments. In that research, prosodic processing advances language in 3 main ways: aiding the segmentation of the speech stream (and, indirectly, the segmentation of wordforms); providing a perceptual basis for proto-lexical categories; and cuing infants into the morphosyntactic structure of their language.

Here, we document a relationship between the perception of prosody and vocabulary size as predicted by its impact on wordform segmentation. In contrast, we find little evidence for either a difference across different word categories (specifically, nouns as compared to predicates and function words, as we predicted would follow from the perceptual bases for proto-lexical categories), or an impact on morphosyntactic development. There at least 3 possible explanations
for this pattern of results. First, it may be that the model in Figure 1 is incorrect: Even if prosody may still be important for online speech perception, perhaps the importance of prosodic processing for lexical and morphosyntactic development in the long term is limited. Alternatively, perhaps the predictor and/or the outcome variables used here were not ideal to capture the relevant constructs. For instance, our prosodic measure involved attention to the envelope differences associated with large prosodic units, not the prosodic differences used to cue function versus content words, or nouns versus verbs; nor those associated with the head direction parameter. Finally, it may be the case that there is more individual variation in infants’ reliance on prosody for segmenting the speech stream, whereas the other two functions of prosody are more stable across children. In fact, as has been argued elsewhere (e.g., Aslin and Newport, 2009), there are many strategies involved in word segmentation, including the utilization of coarticulation, phonotactics, and transitional probabilities, in addition to prosody. It is therefore plausible that individual infants might rely to a greater or lesser extent on some or other of these strategies.

Further research using computer-based learning algorithms and longitudinal designs should contribute to teasing apart those 3 explanations, in addition to shedding light on the relationship between this sensitivity to prosody measure, and other infant predictors of childhood language. With respect to the latter, there is little reason to believe that the prosodic sensitivity measure explored in the present study is co-dependent with measures of native and non-native sounds discrimination (Kuhl et al., 2008; Rivera-Gaxiola et al., 2005). Instead, segmental and prosodic knowledge may complement each other in infants’ encoding and recognition of wordforms, in which case both would independently predict performance in a wordform recognition task in infancy, such as that shown to covary with later language in Newman et al. (2006).

Additionally, most existing work on infant predictors has focused on productive vocabulary, but it may be of great interest to investigate their predictive value with respect to speech processing abilities in toddlerhood. For example, recent work documents strong relationships between speed and accuracy in a lab-based word recognition task at 25 months and measures of vocabulary and language development both before and after this age (e.g., Fernald et al., 2006; Marchman and Fernald, 2008). One may expect that, if prosody is mainly beneficial to the processing of speech, it will correlate mainly with speed and overall accuracy. In contrast, the native/non-native measures should be better predictors of performance when minimal pairs or mispronunciations are involved (Swingley and Aslin, 2000).

Along these same lines, further research is also necessary to pinpoint the dependence of childhood lexicon and morphosyntax from their (hypothetical) infancy analogues. We (as most infant researchers) believe that speech perception experience gained over the first year of life is fundamental to later language acquisition. But is there a true continuity between the representations built during
infancy and those found in childhood? Or is early language acquisition simply about establishing efficient routines for the processing of one’s native language? It would seem that the only way to positively answer these questions is through longitudinal studies assessing the stability in individual differences between e.g. infants’ sound prototypes and children’s sound representations; infants’ wordform lexicons and children’s true (semantic-based) lexicons; and so on.

Nonetheless, regardless of the specific ways in which infant experience may impact child language, the present study shows that 6-month-olds’ sensitivity to clausal prosody predicted their productive vocabulary 18 months later. The difference in vocabulary size found was robust, held across word categories, and occurred within both sexes. Furthermore, it was not explained through basic differences across the two groups in terms of maternal education, infant cognitive skills, or birth weight. Thus, this is the first demonstration of the longitudinal, predictive value of infants’ attention to prosodic well-formedness with respect to later natural language acquisition, which fits in well with the long-standing belief that prosodic processing matters for language development.

References


Gout, Ariel, Anne Christophe, and James L. Morgan. 2004. Phonological phrase bound-
aries constrain lexical access: II. infant data. *Journal of Memory and Language* 51:547–567.


