Cognition 152 (2016) 53–60

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Original Articles

What infants know about the unsaid: Phonological categorization in the absence of auditory input

Céline Ngon [a], Sharon Peperkamp [a,b,*]

[a] Laboratoire de Sciences Cognitives et Psycholinguistique (Département d'Etudes Cognitives, Ecole Normale Supérieure – PSL Research University/Ecole des Hautes Etudes en Sciences Sociales/Centre National de la Recherche Scientifique), Paris, France
[b] Maternité Port-Royal, Université Paris Descartes, Paris, France

ARTICLE INFO

Article history:
Received 10 October 2013
Revised 14 March 2016
Accepted 19 March 2016

Keywords:
Language development
Phonology
Lexical access
Categorization
Output representations
Inner speech

ABSTRACT

Acquiring a lexicon constitutes an essential step in early language development. From an early age on, infants store words with well-specified phonological representations, and they can spontaneously activate these representations on the basis of visual information only (Mani & Plunkett, 2010a, 2011). To what extent can infants inspect and categorize phonological representations in the absence of auditory input? The present study focuses on words that infants comprehend but do not attempt to pronounce yet, and introduces a novel methodology based on anticipatory eye-movements. In two experiments, 21-month-old French-learning infants were silently presented with images of familiar objects whose labels they comprehended but did not pronounce yet. We tested whether they could activate the phonological representation of these labels and categorize them based on their length. Infants' performance exceeded chance when the target words were mono- and trisyllabic, but not when they were mono- and disyllabic. Thus, even in the absence of auditory input infants can activate the phonological representation of words they do not pronounce yet, and use this representation to perform a categorization based on word length, provided the length difference is substantial.

1. Introduction

Building a lexicon constitutes an essential step in early language acquisition. During the first two years of life, infants develop perceptual skills that allow them to recognize word forms, as well as productive capacities for pronouncing those words. Typically, word comprehension starts earlier than word production: whereas infants show some signs of word comprehension as early as 6–9 months of age (Bergelson & Swingley, 2012; Tincoff & Jusczyk, 1999), they only utter their first words around 12 months (Vihman, 1996).

While adding words to their lexicon, infants develop phonological representations for these words. A host of evidence shows that as early as their second year of life, such representations are well-specified (Bailey & Plunkett, 2002; Mani, Coleman, & Plunkett, 2008; Mani & Plunkett, 2007, 2010b; Skoruppa, Mani, Plunkett, Cabrol, & Peperkamp, 2013; Swingley, 2009; Swingley & Aslin, 2000, 2002; White & Morgan, 2008; Zesiger, Dupuis Lozeron, Lévy, & Frauenfelder, 2011). These studies show that infants are sensitive to mispronunciations in both consonants and vowels: they recognize a target object better when its label is correctly pronounced ('Where's the baby?') than when it is mispronounced ('Where's the vaby?'). Thus, infants perceive phonetic details and store them in their phonological representation of words from an early age. In sharp contrast to these detailed input representations stand infants' approximate and highly variable early word productions. Some of this variability is due to systematic alterations, including sound substitution (for example when a target word consonant takes on features of another consonant, a phenomenon referred to as consonant harmony, e.g. *guck* for *duck*; see e.g. Goad, 1997; Vihman, 1978; Pater & Werle, 2003) and word truncation (when a syllable of a word is omitted, e.g. *nana* for *banana*; see e.g. Allen & Hawkins, 1978; Demuth, 1995; Fikkert, 1994; Gerken, 1994; Ingram, 1978; Pater, 1997; Smith, 1973). It is not until around 6 years of age that most words are pronounced correctly (Sander, 1972; Vihman, 1996).

In models of word production, generating a word begins with the selection of a lemma and the retrieval of the associated word form, including its phonological representation. In adults, the activation of phonological representations is rapid and automatic, and it takes place even without the intention to speak. This follows from research with visual search tasks, in which participants

* Corresponding author at: Laboratoire de Sciences Cognitives et Psycholinguistique, Pavillon Jardin, 29 rue d'Ulm, 75005 Paris, France.
E-mail address: sharon.peperkamp@ens.fr (S. Peperkamp).

http://dx.doi.org/10.1016/j.cognition.2016.03.014
0010-0277/

must identify a target object among a set of objects shown on a screen. Crucially, adults are slower to identify an object in the presence of a distractor whose label is homophonous or phonologically similar to the label of the target object. This phonological interference effect can only be due to the automatic activation of the object labels (Görges, Oppermann, Jescheniak, & Schriefers, 2013; Meyer, Belke, Telling, & Humphreys, 2007). There is evidence that infants who have just started to pronounce words likewise activate phonological representations of unnamed objects (Mani & Plunkett, 2010a, 2011). In these experiments, 18- and 24-month-old English-learning infants heard a label and had to recognize the target object to which it referred (presented side-by-side with a distractor object). Crucially, infants were primed with a silent presentation of an object whose label either started with the same phoneme as the label of the target object or a different one (e.g., target: cat; related prime: cup; unrelated prime: teeth). Infants' recognition was found to be significantly different between related and unrelated trials. That is, 18-month-olds looked longer to the target than to the distractor and were faster to switch from the distractor to the target in related than in unrelated trials; 24-month-old showed the reverse pattern of results, and hence behaved similarly to the adults in the studies mentioned above, whose recognition of a target object was also inhibited by the presence of a phonologically related prime (Görges et al., 2013; Meyer et al., 2007). Thus, infants spontaneously activated the phonological representation of the labels of the silently presented prime objects. It remains to be investigated whether this spontaneous activation depends upon infants' capacity to overtly pronounce the word. Indeed, Mani and Plunkett (2010a, 2011) did not consider infants' expressive vocabulary; given that the primes were familiar monosyllables, it is likely that 24-month-olds already pronounced most of them, and that the 18-month-olds produced at least some.

Infants' capacity to spontaneously activate phonological representations raises a further question: to what extent can infants inspect and categorize these representations? To our knowledge, this question has been addressed only with six- and seven-year old children. Specifically, when shown a set of three pictures, six-year-olds can indicate the picture whose label rhymes with an auditorily presented word, and seven-year-olds can also indicate the picture whose label has the largest number of sounds (Lundberg, Frost, & Petersen, 1988; note that the article does not mention whether the children pronounce the labels before giving a response). In the present study, we examine the capacity to categorize phonological representations of unnamed objects at a much younger age. Specifically, using an implicit anticipatory eye-movement paradigm (McMurray & Aslin, 2004), we investigate whether 21-month-old infants, who do not pronounce many words yet, can categorize the labels of familiar objects that are presented in silence according to whether they are short or long. As categorization requires activation, we also test whether infants can activate phonological representations of unnamed objects even if they do not pronounce their labels overtly. Thus, we use words that according to parental report are known but not yet pronounced.

2. Experiment 1

In this experiment, infants have to categorize monosyllabic and trisyllabic labels. Considering the novelty of the paradigm and the difficulty of the combined activation-categorization task, the 1:3 ratio seems an appropriate starting point. This is the highest possible ratio we can test, since French infants do not know a sufficient number of words with more than three syllables.

Table 1
Mean number of segments and mean durations of mono- and trisyllabic words in Experiment 1. SEs are shown between parentheses.

	Number of segments	Duration
Monosyllable	2.9 (0.1)	738 (46.7)
Trisyllable	6.6 (0.2)	966 (42.0)
Difference	*3.7**** *(0.2)*	*228**** *(44.2)*

*** $p < 0.001$ (t-test).

2.1. Methods

2.1.1. Participants

Thirty-one 21-month-old monolingual French-learning infants from Paris participated (9 boys, mean age = 21;17, range = 20;1–22;1). Parental consent was signed prior to testing. Six additional infants were tested but excluded from analysis due to fussiness before reaching the test phase (5) or experimental error (1).

2.1.2. Stimuli

We selected 30 monosyllabic and 30 trisyllabic words representing unambiguously recognizable objects. Both monosyllabic and trisyllabic lists were matched with regard to semantic category: they contained as many objects corresponding to animals (e.g. *chat* 'cat', *papillon* 'butterfly'), food items (e.g. *oeuf* 'egg', *clémentine* 'clementine'), and artifacts (e.g. *lit* 'bed', *parapluie* 'umbrella'). All words were produced in an infant-directed manner by an adult female native speaker of French. Information about mean number of segments and durations is provided in Table 1. In addition, we selected 60 color drawings depicting the objects.

2.1.3. Procedure

We used an anticipatory eye-movement paradigm, consisting of a learning phase and a test phase (McMurray & Aslin, 2004). Personalized scripts were programmed separately for each infant based on their individual comprehension and production vocabularies.

To obtain vocabulary reports, we asked parents to fill out a questionnaire consisting of the selected word list (see Section 2.1.2) which they had to send back a few days prior to test. For each word parents had to indicate whether they thought their child could comprehend the word, pronounced the word and if so, how it was pronounced. For the test phase, we selected images of objects that infants recognized and whose label they comprehended but did not pronounce, according to parental report.[1] Objects for the learning phase were selected among those remaining in the vocabulary list after assignment of the objects in the test phase; that is, they could be either unknown, known and pronounced, or known but not pronounced by the infant.

The trial design was adapted from Kovács and Mehler (2009). During the learning phase, trials began with the central presentation of an image for 1500 ms, while an isolated auditory label for the depicted object was played simultaneously. The offset of the image was followed by two white squares on each side of the monitor for 1000 ms. Next, the same object reappeared within one of the two squares for 1500 ms, again accompanied by its auditory label. For half of the infants, objects with a monosyllabic label always reappeared on the left side of the screen and objects with a trisyllabic label on the right side; for the other half, it was the

[1] The test stimuli for 17 infants included one or more monosyllabic words they already pronounced. In addition, the test stimuli for 13 infants included one or more trisyllabic words that they pronounced in a truncated, monosyllabic, form (e.g., *clémentine* pronounced as *tine*). All trials with a pronounced mono- or trisyllabic word were omitted from the analyses; this concerned 13.7% of the test trials (9.6% with monosyllabic and 4.2% with trisyllabic words).