Computerized test for oral language perception: TIPLO

Christophe Parisse, Edith Soubeyrand

Title: A computerized test for oral language perception: TIPLO

Christophe Parisse and Edith Soubeyrand

Institut National de la Santé et de la Recherche Médicale (INSERM), Paris, France

Running head: computerized test for oral language perception

Keywords : perception, comprehension, phonology, evaluation, young child

Address where correspondence should be send:

Christophe PARISSE

Laboratoire de neuropsychopathologie du langage et de l'enfant

Bâtiment Pharmacie, 3ème étage,

Hôpital de la Salpêtrière

47 Boulevard de l'Hôpital

75651 PARIS CEDEX 13

FRANCE

E-mail address: parisse@ext.jussieu.fr

A computerized test for oral language perception: TIPLO

Abstract : A computerized test for oral language perception (TIPLO) has been developed to evaluate the abilities of young children (from age two) and of children with language disorders (including no language production) to perceive and discriminate phonemic information. The test has a computer-game presentation which makes it easy to use with very young children. It runs on any recent personal computer under PC/Windows. An experiment with 19 normally-developing children aged 18 months to 33 months has shown that children are keen to use the test and that it reveals a good sensibility to the development of phonemic abilities in children from age two to age three. Children with severe language disorders also have no problem with the test.

There are two versions of the test: one with two and one with four images displayed simultaneously on the screen. The one with two images is for very young children, and can also be displayed simultaneously on two screens instead of one. The version with four images is for older children (after 30 months). Children have to point out one of the images displayed after hearing oral instructions given through the computer loud-speakers. The standard test finds out whether a small phonemic difference is perceived (such as a difference in only one vowel or consonant such as ‘sheep’ and ‘ship’) and another version of the test implements a classical designation test with large phonemic differences between items. Each test consists of 20 items and lasts about 15 minutes.

A computerized test for oral language perception: TIPLO

INTRODUCTION

It is often crucial to test the receptive language of young late-talking children, because we know that deficits in language comprehension and deficits in language production give rise to different diagnostics and prognoses. A test that could evaluate the development of language comprehension at the phonemic and lexical levels and could be used with young children or children with severe language delay would therefore be very useful. Such a tool would be aimed at normally-developing children with ages ranging from two to three years and language impaired children with little or no language produced. Because of this, it would have to be very simple, user-friendly and take rather short time.

Bearing all these reasons in mind, we developed TIPLO, a computerized test for oral language perception. The use of the computer had various advantages:

- young children are often interested in and less intimidated by computer or television than by human interaction, and less liable to play with an image on the screen than with a figurine
- no printed material, that children would also be tempted to play with
- verbal instructions are always produced identically
- post-test analysis can be prepared automatically

The first goal of TIPLO was to be attractive to young children, so that the presentation of the multiple choices question is interweaved with short ten-second runs of a cartoon. Making a choice is rewarded by the continuation of the cartoon. The second goal of TIPLO is to test the phonemic abilities of children. Choices include picking a picture among distractors that contain a word with very close phonetics. The instructions and questions are delivered by the computer which provides an identical phrasing and pronunciation for all test runs. As TIPLO targets young children –children whose receptive language would correspond to the

two-word stage or just later–, the items selected to test their phonemic abilities are all words that are acquired very young in normally-developing children. The items cover the perception of consonants, vowels, clusters and functional words.

Choice of the protocol

When testing is done with very young children, it is often necessary to use a technique that does not require active participation from the child such as the intermodal preferential looking paradigm (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987). Other techniques such as the high-amplitude sucking technique and the visually reinforced head-turn cannot be used with children aged eighteen months and more. The problem with the intermodal preferential-looking paradigm is that it takes a long time and often demands a lot of supplementary work when computing the results. This is due to the fact that the side which the child looks towards preferentially is not always clear and it becomes necessary to do a lot of post-testing verifications using video-recordings with inter-rater reliability. All this is time consuming and makes taping the test procedure mandatory, which calls for material and requires special authorization from the child's parents.

The procedure most often used when testing with older subjects is the picture selection task (e.g. Barton, 1976; Barton, 1980). Here, post-testing analysis is much more simple as the child produces a conscious and explicit decision, usually completely unambiguous; in case of doubt, a decision is taken on-line by the practitioner or the item skipped. The drawback is that this type of test requires a comprehension of the procedure, often given orally by the practitioner. One particular problem is that one cannot know whether a child with severe language development problems understands the task that is required of her. Also, young children or children with severe language deficits tend to find the procedure boring or strenuous because of it is difficult for them. It is thus necessary in such cases to find some way of motivating the children into following the procedure to its end.

Our goal is to make the picture selection task usable with very young children in a user-friendly way and to obtain a test that gives a quick and easy measure of the quality of the language comprehension and the phonemic perception abilities of a child.

Choice of the phonetic elements to be tested

It is impossible to test all the possible phonemic knowledge of a young child in a single test, as it is necessary to limit its duration to about a quarter of an hour. A careful look through the literature on the development of phonemic knowledge should give us the most interesting and informative phonetic contrasts to test at a given age. It is also necessary to find which words can be used at what age to test the children's knowledge. These words have to be known by most children because we want to use the same list of words with every child.

It was impossible to limit ourselves to the use of data from studies in language perception to find out which phonemic patterns were the most interesting to study with TIPLO. This is due to the scarcity of such studies after age one, for the following reasons. First, perception studies are more complex when children are young and cannot follow the protocol of classical psychological experimentations. Second, it is often believed that comprehension precedes production and that it is thus not necessary to test children for knowledge in comprehension when they display this knowledge in production. Phonemes produced correctly and consistently will usually not be tested in perception, even though this correct production may result from holistic or partial knowledge. So, we will first present data from studies in perception, and then data from studies in production.

Perception studies begin with the birth of the child. Categorial discrimination of phonemes can be found in children a few days old (Eimas, 1974; Bertoncini, Bijeljac-Babic, Blumstein, & Mehler, 1987). Before six months of age, children were shown to discriminate many minimal pairs of phonemes. Extensive bibliography about the subject can be found in Jucszyk (1997) and Vihman (1996). For instance, discriminations were found in various

languages between pairs such as [bæ] and [dæ], [wa] and [ja], [fa] and [θa], [va] and [ða], [ra] and [la], [ba] and [wa], [s] and [z], (Li & Ma,) and [ba], [pa] and [pã], [ʊ] and [Y], [u] and [y], [a] and [i], [i] and [u], [i] and [ɪ], [a] and [ɔ], and [a] and [i]. This impressive list shows that many phonetic contrasts can be recognized by young children, but should not hide the fact that many other contrasts were never tested and that the results obtained are only statistical tendencies often limited to CV contexts. When phonemes are inserted in more complex contexts, as shown by Goodsitt, Morgan, and Kuhl (1993), 6-month-olds have some problem differentiating contrasts such as [kotiba] and [kotidu]. As we said, works about the development of perceptual abilities after age one as less numerous. Shvachkin (1973) studied the development of perception in eighteen Russian children from age 0;10 to age 2;0. Children were trained with non-words with a VC or CVC format. These word formed minimal pairs such as [mak] and [rak]. These pairs were later used to test the children's perceptual capacities. Shvachkin found twelve stages of development (see Table 1). The progression from one stage to another follows Jackobson's theory of phonetic development. Garnica (1973) replicated the study with American children and found similar results, but with more variability from one child to the other, so that Shvachkin's twelve stages should be taken more as a general pattern to be used in research than a precise guide to the development of phonemes perception.

insert Table 1 about here

Criticism about Shvachkin's study bear on the fact that he used CVC words, which makes phonemes easier to recognize. His claim that perception development ends at age two is thus not generally accepted. Barton (1980) also cast doubts on Shvachkin's and Garnica's statistical procedures and upon the validity of Garnica's ordering of the acquisition of English

phonemic discriminations. More complex syllables do seem to give more problems. For example, Edwards (1974) in a similar study done with older children – 1;8 to 3;11– found that "phonemic perception develops in a gradual and patterned way". Barton (1976), although working with yet older children – age 2;3 to 2;11 – obtained his weakest results in items which contained clusters.

Data about the development of production is more detailed. Sander (1972) found that phonemes /p/, /m/, /h/, /n/, /w/, /b/ start to be produced by fifty percent of English children at age 1;6, phonemes /k/, /g/, /d/, /t/, /ŋ/ at age 2;0, /f/, /j/ at age 2;6, /r/, /l/, /s/ at age 3;0, /tʃ/, /ʃ/, /z/ at age 3;6, /dʒ/, /v/ at age 4, and other consonants later on. Stoel-Gammon (1985) found that phonemes /b/, /t/, /d/, /k/, /g/, /m/, /n/, /h/, /j/, /w/, /f/, /s/ are produced in initial position by 50% of English children at age 2;0. In final position, she found only phonemes /p/, /t/, /k/, /m/, /n/, /ŋ/, /r/, /s/. For French children, Vinter (2001) found a production of /p/, /b/, /t/, /d/, /k/, /m/, /n/, /f/, /v/, /s/, /l/, /w/, /j/ in initial position, while the only phonemes produced in final position were /p/, /t/, /k/, /m/, /n/, /l/ and /R/.

Stoel-Gammon and Herrington (1990) in a review of the literature report that the American English vowel system is generally completed before the consonantal system is complete. They quoted results that showed that many vowels are acquired at age 1;8, but the largest quoted studies (Wellman, Case, Mengert, & Bradbury, 1931; Hare, 1983) showed that vowels /i/, /u/, /ɑ/, /o/, /ʌ/, and /ɘ/ are mastery by 75% of the children at age two, with some vowels above 90% correct and the other vowels being mastered later. Stoel-Gammon (1991), quoting Templin (1957), indicate that, at three years, normative data would indicate that 93.3% of the vowel phonemes are produced correctly on average.

Choice of the items

Many phonetic studies carried out with young children first ensure that all test items are known to the child to be tested. Studies such as Shvachkin's go even further and all the children are trained with nonce items. However, as this procedure takes a lot of time, we did not used non-words nor tested whether children knew the test words beforehand, in order to maintain the duration of the whole testing procedure under a quarter of an hour.

To balance this out, we tried to ensure that the item tested were known by most children under age two or age three by using results from the Macarthur Communicative Development Inventories (Fenson et al., 1994) and from several studies and tests about the development of French children (Chevrie-Muller, M., & Decante, 1981; Chevrie-Muller, Simon, Le Normand, & Fournier, 1988; Arabia-Guidet, Chevrie-Muller, & Louis, 2000). The list is unfortunately not very large as we also needed to find out word pairs that presented a minimal phonetic difference. Furthermore, the production of a word by children does not prove that they are able to match picture and word easily. This is especially true for abstract words such as 'cold', 'good', and colors. So, we tested a first list of twenty-five couples of words, out of which only eighteen items were finally selected. The initial testing and results which lead us to the final version of the test is presented below.

FORMAT OF THE TEST

The choice of the format of the test intents to fulfill four goals:

1. easy and quick to use for both child and practitioner
2. usable with children with very low language comprehension level or with attention problems
3. can be used on a low price computer
4. has possibilities of further extensions

For an easy exploitation by the practitioner, the principle of the picture selection task has been retained. Children are asked to point at a target image to be chosen from a selection of two or four images. There are two versions of the test where the child is presented with two images, and two versions where the child presented with four images. In each case, one version presents material with minimal phonetic contrast between target word and main distractor – for example ‘sheep’ and ‘ship’ for two images, or ‘sheep’, ‘ship’, ‘cow’ and ‘house’ for four images – and the other version presents material with maximal phonetic contrast between the target word and the distractors, which amounts to a simple designation task. This second version is used to control whether children do know the words illustrated, and whether the presence of a small phonetic contrast makes the task more difficult or has no effect.

Target images were selected so that the less attractive images are the target images, otherwise children would give the correct answer all the time. After our initial tests, images from our material that turned out to be particularly attractive to children were mostly those representing babies, cars, balloons, dogs, and cats. None of these words are target words, with the exception of ‘babies’ which is contrasted with the singular ‘baby’. In this case, the singular form appears to be the ‘default’ choice for the children, so our target is the plural form.

The use of a computer allows pre-recorded speech signals, which guaranties that every children tested will have the same input, and an automatic recording of the results. The test runs on a PC under Windows 98 with a processor of 200 MHz and 32 Mo of memory. It is possible to use additional hardware features such as a touch screen or two screens instead of only one, but all test features can be used without any additional hardware.

Computer are easy to use with young children because they are often used to television screens or even real computers and less prone to playing with the screen than with images or

small toys. To make the test more attractive to young children, short strips of a cartoon are inserted between the presentation of the test images. The child is asked to choose an image on the screen, before being allowed to see the next part of the cartoon – the detailed description of the test protocol is given in Figure 1. This cartoon helps in keeping young children interested until the end of the test.

insert Figure 1 about here

Choice of the test items

insert Table 2 about here

The complete list of items selected in the final version of the test is presented in Table 2. Two supplementary items are set at the beginning of the test to help the child use the system. The test items presented in Table 2 correspond to pairs with a minimal phonemic distance. For the test with maximal phonetic distance, the target item is unchanged and the distractor replaced by a word very dissimilar to the target word. All pairs do not correspond to a single difference in phoneme because we were limited to words that most young children will know, so that only fourteen contrasts do so, such as for example, 'grand (tall)' [gRã] vs. 'gros (large)' [gRo]. In six cases, the difference bear on more than one phoneme – for example, 'bras (arm)' [bRa] vs. 'froid (cold)' [fRwa], or 'ventre (belly)' [vãtR] vs. 'chambre (room)' [ʃãbR]. In all cases except two, we have the same syllabic structure. The two exceptions are CVCC vs. CVC ([ʃãbR] vs. [ʒãb]) and CCV vs. CCvV ([bRa] vs. [fRwa]).

Fourteen test words consist of only one syllable. There are six items with CV structure, three with CVC structure (including the CVCC exception), one with CvVC structure, and four

with CCV structure (including the CCvV exception). Three test items have two syllables, all with CVCV structure. One test item contains three syllables, with a CVCVCV structure. The range of syllable structures presented allows to test whether recognition is better in monosyllabic or polysyllabic words.

Eighteen different phonetic contrasts are tested. It is not possible to test all types of phonetic contrasts with our vocabulary limits, but we strived to reach a relative balance between consonantal contrasts, vowel contrasts, and contrasts within clusters, as well as to have the largest possible variety of contrasts. Contrasts involve voiced vs. non-voiced, nasal vs. oral, manner of articulation, place of articulation, open vs. close vowels, front vs. back vowels. About half the contrasts involve a single phonetic feature, and the other half more than one. Results should be better with large phonetic contrasts, and with those involving stops or vowels.

Significance of the results

For the version of the test with two pictures, results are significant if children give fourteen good answers out of twenty, $\chi^2(1, N = 18) = 5.55$, $p = 0.019$. For the four picture test, the number of good answers required is ten, $\chi^2(3, N = 18) = 8.96$, $p = 0.029$, and significant scores can also be computed for 'consonants only' (6 good answers out of 8 required, $\chi^2(3, N = 18) = 10.66$, $p = 0.013$), and for 'vowels only' (5 good answers out of 6 required, $\chi^2(3, N = 18) = 10.88$, $p = 0.012$).

INITIAL TESTING AND RESULTS

Subjects

Tests were conducted with nineteen subjects from a nursery of the Hôpital de la Salpêtrière. All children had normal cognitive and language development. The children's age ranged from 1;6 to 2;9. Some children ran through the test a couple of times, either within a

very short delay (one week) or else after a long one (three months). This made a total of thirty observations. Ages of testing for all children are presented in Table 3. The results are presented in three parts, by age groups – before age 2;0, age 2;0 to 2;3, after age 2;6 –, because the age factor appears to be very important in determining what type of results are obtained and in which way TIPLO can be most useful.

insert Table 3 about here

Results for subjects before age 2;0

Very young children – about age 1;6 – have many problems with the test. They find difficult to understand what they are asked to do. It is in this age group that we have our sole 'no answer at all' results. The administration of the test at that age often requires a long preparation of the child which makes it unsuitable for systematic use. Even in the case of the two children who did understand the procedure, the quantity of responses were not very good, with answers to only half the items. However, when they did answer, they did it rather well: their few answers did not appear to be at random.

With slightly older children – about age 1;9 –, the results were much better, with only one child out of four who did not understand the purpose of the test quickly, and even this child eventually gave correct answers. At that age, pointing at a picture with hand or finger is not always something that comes naturally to the child. At this age, the use of two screens becomes mandatory because the large size of the images helps to determine with some accuracy which picture was chosen by the child.

One of the major surprise of this test phase was the fact that all children before age 2;0, and some children up to age 2;3 often refused to give an answer, even when they perfectly understood the purpose of the test. Even when further prompted, the children were still reluctant to give 'some answer' even though they did take the time to look at the different

pictures, and when they finally accepted to answer, it was mostly by pointing to some part of the screen irrelevant to the question asked, but familiar to them. As if, to a certain degree, they knew that they did not, in fact, know.

The systematic aspect of this behavior lead us to compute two types of score: one with reference to each questions answered to by the child (called below relative score); and one with reference to all the questions, with or without answers (called below absolute score) – see column 4 of Table 3. Whereas there was only one significant absolute score before age 2;0, they were four significant relative scores in the same age group. This shows that it is possible to use TIPLO before age 2;0, but the small number of significant results and the difficulties met while administrating the test limit its usefulness, at that age, to fundamental research investigations.

Results for subjects at age 2;0 to 2;3

After age 2;0, it becomes much easier to use TIPLO. Most children understand immediately what is the purpose of the game, and they all like to play with the computer. However, their results are not yet very good. Most children had significant relative scores – 6 out of 9 –, but only one – 1 out of 9 – has a significant absolute score. This would mean that the comprehension vocabulary is not yet large enough to allow them to understand all the items. There is also a clear difference at that age between the two-picture and the four-picture version of the test. Although children are able to use the four pictures version, it is clearly much more difficult for them, it takes a much longer time to go through the test and the results are not very good. It seems therefore necessary to confine ourselves to the two-picture version, which limits its sensitivity as testing the perception of some specific contrasts such as between consonants or between vowels cannot be done without running the test several times.

Results for subjects after age 2;3

All children used the test without any major problem. All but one of the eleven children aged more than 2;3 had both training items correct, and the child who had only one training item correct did very well on the rest of the test (child 19 in Table 3). It was sometimes, especially for the youngest of these children, necessary to repeat some items of the test as children's attention did wander at times, but this was less necessary than with children aged less than 2;3. The number of twenty items is adequate as it makes the whole test last about a quarter of an hour, which seems a good length duration for children at that age. Although all children used the test easily, not all children reached a significant result. For children aged 2;3 to 2;6, there is still a difference between relative and absolute results.

After age 2;6, children appear to be a lot less inhibited and answer to all items. This could also be an effect of their increasing lexical knowledge. It is difficult to know whether errors are due more to insufficient phonetic performance than to a lack in lexical knowledge because we have tested too few children as yet. In the first results (for four children aged 2;6 or older), there appears to be two types of errors (phonetic and semantic), as well as some unclassifiable errors. The types of phonetic errors cannot be reliably analyzed with the current number of children that passed the test. Children did not reach maximum scores, which means that it is necessary to continue to test older children to determine when the test stop being useful.

Six children (#5, 7, 10, 12, 17, and 18) ran through the test twice within a very short length of time (either just after the first time, or one week later). Not all situations were equal because some children were tested using exactly the same conditions –number of pictures and test complexity– and others were not – they had a different number of pictures or a different test complexity –, but no facilitation effect was put in evidence. There is a clear difference between the results obtained for the two applications of the test for one child, but this was a

young child and filling a bit ill during the second time around. In all other cases, both results were quite similar. Four children were tested twice with a long delay between the tests – three months. In all cases, results were better the second time, which demonstrates the developmental sensitivity of the test.

CONCLUSION AND PERSPECTIVES

One of the eternal questions in research about early language development is to know what children really understand, and how they use this knowledge to develop and produce language. There are more studies in language production than in language perception for the obvious technical reason that it is easier to gather and analyze produced language than perceived language. However, when a child begins to speak, she produces only approximations of the language she hears. Is this a limit of her perception or of her production? TIPLO can help to target such interesting issues, and other issues such as whether young children's perception is holistic or analytic.

Of course the same questions apply to children with language development difficulties, which emphasizes the importance of adequate tools to evaluate language perception. Although TIPLO can certainly be useful to study the normal development of language, it was first intended to be used to study and evaluate perception and lexical comprehension in children with language impairment, and especially young children with language impairment. It is very important to be able to evaluate the verbal comprehension of children that often cannot yet produce language for an early diagnostic of language impairment. Although it cannot be done quickly and easily, it is necessary, because a deficit in comprehension changes diagnostics, prognoses, and remediation strategies. TIPLO has already been used with children with severe language impairments in our laboratory, and these children have run the test as easily as younger normally-developing children. This is very important as it is often

very difficult with these children to determine if their problems with usual tests come from language impairment, attention problems, or their plain good will, or lack of it.

The number of normally-developing children tested until now is too small to demonstrate the reliability of the test in a statistically meaningful way. However, tests performed thirty times with nineteen children at different ages showed evidence that it fulfills its main purpose quite well, as quick tests of the 'picture selection' type can easily be used with children under three. TIPLO cannot be used to test every phonetic contrast, but this can be done later on by a speech pathologist if necessary. As it stands, no long preparation is necessary to determine which items are known to the child and no long post-processing is necessary to compute the results, which is very important for clinical settings.

TIPLO can already be used in its actual form and can be freely provided for research and medical purposes by the first author of this article. However, further developments can be made to it. First, more testing have to be done with children aged 2;6 to obtain valid norms. It is also necessary to test children older than 3;0 to know the maximal values the scores can reach. Second, it would be interesting to use this test as a starting point for developing future more complex tests using the same tool but targeting consonants, vowels, clusters, or functional words specifically.

REFERENCES

Arabia-Guidet, C., Chevrie-Muller, C., & Louis, M. (2000). Fréquence d'occurrence des mots dans les livres d'enfants de 3 à 5 ans. Revue européenne de psychologie appliquée, 1, 3-16.

Barton, D. (1976). The role of perception in the acquisition of phonology. Unpublished Doctoral dissertation, University of London.

Barton, D. (1980). Phonemic perception in children. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), Child phonology 2: Perception . New York, NY: Academic Press.

Bertoncini, J., Bijeljac-Babic, R., Blumstein, S. E., & Mehler, J. (1987). Discrimination in noenate of very short CVs. Journal of the Acoustical Society of America, 82, 31-37.

Chevrie-Muller, C., M., S. A., & Decante, P. (1981). Epreuves pour l'examen du langage. Paris: Editions du Centre de Psychologie Appliquée (ECPA).

Chevrie-Muller, C., Simon, A. M., Le Normand, M. T., & Fournier, S. (1988). Batterie d'évaluation psycholinguistique - Révision du manuel 1997. Paris: Edition du centre de psychologie appliquée (ECPA).

Edwards, M. L. (1974). Perception and production in child phonology: The testing of four hypotheses. Journal of Child language, 1, 205-219.

Eimas, P. D. (1974). Auditory and linguistic processing of cues for place of articulation by infants. Perception and Psychophysics, 16, 513-521.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D., & Pethick, S. J. (1994). Variability in early communicative development. Monographs of the Society for Research in Child Development, 59(5), 1-173.

Garnica, O. K. (1973). The development of phonemic speech perception. In T. E. Moore (Ed.), Cognitive development and the acquisition of language . New York, NY: Academic Press.

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.

Goodsitt, J. V., Morgan, J. L., & Kuhl, P. K. (1993). Perceptual strategies in prelingual speechs segmentation. Journal of Child Language.

Hare, G. (1983). Development at 2 years. In J. V. Irwin & S. P. Wong (Eds.), Phonological development in children: 18-72 months . Carbondale, IL: South Illinois University Press.

Jucszyk, P. W. (1997). The discovery of spoken language. Cambridge: MA: MIT Press.

Li, W., & Ma, M. (1992). Ertong Lijie Suliangci 'jige', 'henduo' de Fazhan Tedian (A study on the developmental characteristics of children's capacity in understanding numeral-classifier compounds 'some' and 'many'. Acta Psychologica Sinica, 1992, no. 2, 158-164.

Sander, E. (1972). When are speech sounds learned? Journal of Speech and Hearing Disorders, 37, 55-63.

Shvachkin, N. K. (1973). The development of phonemic speech perception in early childhood. In C.-A. Ferguson & D. I. Slobin (Eds.), Studies of child language development (pp. 91-127). New York: Holt.

Stoel-Gammon, C. (1985). Phonetic inventories 15-24 months - a longitudinal study. Journal of Speech and Hearing Research, 28, 505-512.

Stoel-Gammon, C. (1991). Normal and disordered phonology in two-year olds. Topics in language disorders, 21-32.

Stoel-Gammon, C., & Herrington, P. (1990). Vowel systems of normally and phonologically disordered children. Clinical Linguistics and Phonetics, 4, 145-160.

Templin, M. C. (1957). Certain language skills in children: Their development and interrelationships. (Vol. 26). Minneapolis, MN: University of Minnesota Press.

Vihman, M. M. (1996). Phonological development. Oxford: Blackwell.

Vinter, S. (2001). Les habiletés phonologiques de l'enfant de deux ans. Glossa, 77, 4-19.

Wellman, B. L., Case, I. M., Mengert, J. B., & Bradbury, D. E. (1931). Speech sounds of young children. University of Iowa Studies in Child Welfare, 2(5).

Figure 1 : Description of a test sequence

There are twenty test sequences in the whole test. Each sequence follows the same steps.

Step 1:

A cartoon sequence – 10 to 15 seconds – is presented on the screen (the same sequence on each screen for the presentation using two screens).

Step 2:

The cartoon freezes and a pre-recorded voice tells the child: ‘If you want to know the rest of the story, then show me the bath (for example).

Step 3:

Two images appear side by side (both on the same screen or one per screen for the presentation using two screens). For presentation with four images, they all appear on the same screen.


Step 4:

The child has to choose one of the images. She is invited by the practitioner to stand up and touch the screen, if she wants to. Or, she may just point towards the image with her hand. If the child does not answer before ten seconds, the end of the instruction is repeated: ‘Show me the bath’.

Step 5:

When the child has chosen an image, she goes back to her seat and the practitioner starts the next sequence by hitting the key that corresponds to the child’s answer:

- left-arrow for left image
- right arrow for right image
- space bar for no answer or unclear answer

This allows the computer to store the data, and go back to step 1.

Table 1: Synopsis of Shvachkin's twelve stages of phonetic perception development

Stage	Difference learned in the current stage	Example of contrasts
1	Between vowels	mak-mek
2	Presence and absence of an initial consonant	ak-tak
3	Sonorants (m, n, r, l) and stops (articulated obstruents)	bak-mak, nak-rak
4	Palatalized and non-palatalized	pʲak-pak
5	Between sonorants	mak-nak, nak-lak
6	Sonorants and fricatives (non-articulated obstruents)	mak-zak, lak-xak
7	Labials and linguals	bak-dak, bak-gak
8	Stops and fricatives (spirants)	bak-vak, kak-xak
9	Pre- and post-linguals	dak-gak, sak-xak
10	Voiced and voiceless consonants	pak-bak, fak-vak
11	Dentals and alveolars (hushing and hissing)	ʃak-sak, ʒak-zak
12	Liquids and semi-vowels	rak-jak, lak-jak

Table 2: Items of the test.

	# item	Phonetics	Target item	Phonetic distractor
Consonants	1	[ʃ] vs. [ʒ]	Jambe (leg)	Chambre (room)
	2	[s] vs. [n]	Soir (evening)	Noir (black)
	3	[R] vs. [l]	Lit (read)	Rit (laugh)
	4	[ʒ] vs. [R]	Joue (cheek)	Roue (wheel)
	5	[m] vs. [v]	Verre (glass)	Mer (sea)
	6	[R] vs. [m]	Ranger (clean)	Manger (eat)
	7	[b] vs. [p]	Poupée (doll)	Bébé (baby)
	8	[m] vs. [b]	Bain (bath)	Main (hand)
Vowels	9	[ə] vs. [o]	Cheveux (hair)	Chevaux (horses)
	10	[ã] vs. [o]	Gros (tall)	Grand (large)
	11	[o] vs. [a]	Chaud (hot)	Chat (cat)
	12	[o] vs. [ã]	Dent (teeth)	Dos (back)
	13	[o] vs. [õ]	Beau (nice)	Bon (good)
	14	[e] vs. [y]	Nu (nude)	Nez (nose)
Clusters	15	[gR] vs. [pR]	Grand (big)	Prend (take)
	16	[bR] vs. [fR]	Froid (cold)	Bras (arm)
	17	[bR] vs. [tR]	Chambre (room)	Ventre (belly)
Functional words	18	[ə] vs. [e]	Les bébés (babies)	Le bébé (baby)

Table 3: Full results of TIPLO for 19 children aged 17 to 33 months (30 observations)

# child	Age at test	Type of test	Score	Score for training items	Significance of absolute score	Significance of relative score	Notes
1	1.5;08	2S	**0. 0. 9**	0			no answer at all
2	1.5;10	2S	**7.10.18**	1			
3	1.6;17	2S	**0. 0.18**	0			goal of the task not understood
4	1.8;24	2S	**9.12.18**	1		0,08	
5	1.9;03	2S	**6. 6.18**	2		*	appear to be difficult for the child
6	1.9;12	2S	**8. 9.18**	1		*	
5	1.9;14	2S	**1. 5.18**	0			appear to be difficult for the child
7	1.9;16	2S	**14.14.18**	1	*	***	
7	1.9;25	2D	**4. 8.18**	0			child was slightly ill
8	1.9;29	2S	**9. 9.18**	1		**	
9	2.0;04	2S	**13.15.18**	2	0,06	**	
10	2.0;09	2D	**11.13.18**	2		*	
10	2.0;18	2S	**11.14.18**	2		*	
11	2.1;09	2D	**8.12.18**	0			shy child
12	2.1;15	2S	**5. 6.18**	1		0,10	shy child
13	2.1;28	4D	**4. 8.11**	2			stopped before the end
14	2.2;00	4S	**11.17.18**	2	**	**	
12	2.2;07	4D	**6. 8.17**	1		*	shy child
15	2.2;09	4D	**6.15.17**	2			

16	2.3;24	4D	**9.15.17**	2	0,07	*	
17	2.4;15	2S	**14.18.18**	2	*	*	
17	2.4;15	4D	**10.14.17**	2	*	**	
18	2.4;24	4D	**8.17.18**	2			
18	2.5;02	4S	**10.14.18**	2	*	**	
15	2.5;17	4D	**10.17.18**	2	*	*	
19	2.5;30	4D	**7.16.18**	2			
16	2.7;06	4D	**11.18.18**	2	**	**	
18	2.8;07	4D	**12.18.18**	2	***	***	
17	2.8;19	4D	**15.18.18**	2	***	***	
19	2.9;02	4D	**13.17.18**	1	***	***	

Notes:

The figures in the 'score' column correspond respectively to:

- The number of correct answers
- The number of items answered
- The total number of item tested, including the items with a 'no answer' response (if this number is smaller than 18, it means that the test was interrupted, or that the result for an item had to be ignored for technical reasons).

'*' means that significance level is p = 0,05, '**' p = 0,01, and '***' p = 0,001. Marginally significant values are given explicitly. A relative score is computed relatively to the number of explicit responses (no response is ignored). Absolute scores take all responses into account.

2S corresponds to the test with two images and great phonetic contrasts

2D corresponds to the test with two images and small phonetic contrasts

4S corresponds to the test with four images and great phonetic contrasts

4D corresponds to the test with four images and small phonetic contrasts

Specific cases, child per child:

- Child #7 (2nd test) and #11: no answer for the first six items.
- Child #13 : the test was stopped because the child was not paying attention anymore.
- Child #12 : very shy child, so that it was difficult to know which screen she pointed to.
- Child #16 : bilingual child (French and Swiss-German).
- Child #17(2nd test): test realized just after the first one.

Acknowledgments

This work was supported by a grant from INSERM, France: Contrat de Recherche Inserm (4U009B). We would like to thank all the people at the Hôpital de la Salpétrière who made this possible and specially the nurseries' staff. Grateful thanks to Anne Reymond for proof-reading this text.