# Using cluster analysis to explore how children use semantic cues when listening to sentences presented with a babble noise

J. Lagacé, C. Giguère, S. Breau Godwin

University of Ottawa, 451 Smyth Rd., Ottawa, ON, Canada, K1H 8M5, Josee.Lagace@uottawa.ca, Christian.Giguere@uottawa.ca and Stephanie.Breau@uottawa.ca

## INTRODUCTION

It is well documented that children require a higher signal-to-noise ratio (SNR) than adults to understand speech in the presence of background noise (Fallon et al. 2002; Picard & Bradley 2001). This difference in speech perception abilities in noise between children and adults has been attributed to developmental factors including, but not limited to, immature auditory system (Fallon et al 2000, 2002; Talarico et al. 2007), language skills (Elliott et al. 1979) and other cognitive functions, such as short term memory (Choi et al. 2008). Because of these developmental factors, variability on speech in noise tests can be expected both within and across age groups. Indeed, according to Leibold & Neff (2007), many studies on speech perception in noise with normal hearing children report large variability. Adult-like performance on perception in noise tasks is generally not reached until 13-15 years of age (Crandell & Smaldino 2000).

The ability to benefit from the linguistic and contextual information counts among the factors that have been studied to explain the difference between children and adult performance while listening to speech in acoustically challenging conditions (Fallon et al. 2002; Elliott et al. 1979). There are indications that the ability to use semantic contextual cues emerges from as young as 19 months (Friedrich & Friederici 2004). However, the contribution of linguistic contextual cues to word recognition in noise is not clear. Fallon et al. (2002) reported that groups of young listeners of 5 and 9 years of age, as well as adults, show comparable gains from linguistic context when listening to speech in noise. On the other hand, Elliott et al. (1979) found that the performance of children of 9 years of age was significantly poorer than that of children of 11 years of age on the highly predictable sentences of the Speech Perception in Noise Test (SPIN) (Kalikow et al. 1977) at a SNR of 0 dB. Further clarification on the developmental changes regarding the ability to benefit from linguistic context is an important issue to resolve since children are often expected to understand and learn from speech presented in relatively noisy classrooms (Picard & Bradley 2001).

There are indications that normal hearing children show individual variability in their ability to benefit from linguistic context when listening to speech in noise. One underlying factor may be the acquisition of linguistic skills, which is not completed until adolescence and show inter individual differences (Bloom & Lahey 1978). Personal factors, such as living in a bilingual family environment, may also affect the acquisition of linguistic skills or vocabulary and induce differences in the ability to benefit from contextual cues for understanding speech in noisy backgrounds (e.g. Mayo et al. 1997). Few studies have explored individual variability in the amount of benefit derived from the linguistic context (e.g., Rönnberg et al. 1998; Grant & Seitz 2000) and, to the knowledge of the authors, none of these studies have explored this variability among children groups.

Because of the expected variability in children performance, group analysis may not always be the best strategy to explore the benefit from linguistic context since averages can be affected by very high and low scores. An alternative is to conduct cluster analysis to determine the presence of specific profiles within a group. Cluster analysis is a statistical tool used to classify data into subgroups, and does not require any prior knowledge of which element will affect the grouping or which groups will be formed (Kacprzyk & Wesam Ashour Wu 2009). In a cluster analysis, every interdependent relationship is considered and no distinction is made between dependent and independent variables.

The objective of the present study is to explore normal hearing children's variability in benefiting from linguistic context when listening to speech in background noise. For this, the *Test de Phrases dans le Bruit* (TPB) (Lagacé et al. 2011) will be used. The TPB is a Canadian French speech-in-noise test that was recently developed following an approach similar to the one used for the development of the original English version of the SPIN test. The TPB includes five recorded lists of forty Canadian French sentences. Like in the SPIN test, in each list, half of the sentences are highly predictable (HP), which means that identification of the key word can be facilitated by contextual information from the sentence (e.g. *The candle flame melted the* <u>*wax*</u>; Kalikow et al. 1977). The other half comprises 20 sentences with low predictability (LP), conveying little contextual information (e.g. *Paul can't discuss the* <u>*wax*</u>; Kalikow et al. 1977). The listener is required to repeat the final word (key word) of each sentence. Performance is measured as the percent correct score on HP and LP sentences separately in each list. The difference in key word recognition between the HP and LP sentences gives information about the listener's ability to benefit from contextual cues. The TPB is intended to be used with both school-aged children and adults. The data presented in this study were gathered during the normative data collection for the TPB.

## METHODS

### Participants

Prior to the commencement of the study, ethics approval was obtained from all the institutions where the research was conducted. A sample of 69 participants (46 females and 23 males) aged between 9 years and 2 months and 12 years and 5 months, all native Canadian French speakers, was recruited from French schools in the Ottawa region (Ontario, Canada). One parent of each participant signed the consent form and completed a questionnaire to rule out the presence of any exclusion criteria.

### Procedure

Each participant was tested individually in a quiet room in the school with ambient noise levels not exceeding specifications for hearing screening in schools (ASHA 1997). A hearing screening at a hearing level of 20 dB HL was performed with a portable audiometer (Madsen Midimate 622) and TDH-39 headphones prior to the experimental phase. All the participants had normal hearing sensitivity at 500, 1,000, 2,000 and 4,000 Hz, bilaterally. Participants were told that they would hear sentences presented first with babble noise and then without noise. They were asked to repeat the last word of each sentence and encouraged to guess if necessary. Sentences from the TPB were used as well as eight-talker (4 males and 4 females) European French babble (Perrin & Grimault 2005). This pre-recorded European French babble was chosen over recordings of English babble because it better approximated the type of masking experienced by the listeners in daily life (no standardized recording of speech babble by Canadian French speakers was available at the time of the study). The TPB sentences and the babble noise were presented diotically via headphones.

Participants first listened to a practice list of 10 sentences presented at a SNR of +4 dB. The experimental phase then consisted of 3 other lists presented at different SNRs. Total experimentation time was approximately 20 minutes. Children were then invited to choose an item (hockey cards, stickers, notepad, pencils, etc.) as a small thank you gift for their participation.

## RESULTS

### Effect of age

A subgroup (n=10) of children (6 females and 4 males) aged from 9 years 8 months to 10 years 11 months (the 9-10 years old group) and another subgroup (n=11) of children (9 females and 2 males) aged from 11 years 2 months to 12 years 5 months (the 11-12 years old group) were administered the TPB to explore the effects of age on performance in interaction with the amount of contextual information and the degree of difficulty. The TPB was administered at three SNRs: -2, 0 and +2 dB. The order of presentation of the SNR was different for each participant.

Performance scores for the two age groups (9-10 and 11-12 years old) are given in Table 1. The average performance for each type of sentence and the difference score were slightly higher for the 11-12 age group than the 9-10 age group. A mixed ANOVA was conducted with the factors *SNR* (3 levels: SNR -2, 0 and +2), *Sentence type* (3 levels: HP, LP, and difference score) and *Age* (2 levels: 9-10 and 11-12 years of age). *SNR* and *Sentence type* were repeated measures variables while *Age* was an inter-subject variable. Results from the ANOVA revealed a significant effect of the main factors *Sentence type* [$F_{(2,38)}$ = 192.63, $p$=0.000, $\eta^2$=0.91)] and *SNR* [$F_{(2,38)}$ = 182.60, $p$=0.000, $\eta^2$=0.91)], but not *Age* [$F_{(1,19)}$ = 1.90, $p$=0.184, $\eta^2$=0.09]. The double interactions *Sentence type* X *Age* [$F_{(2,38)}$ = .20, $p$=0.842, $\eta^2$=0.01] and *SNR* X *Age* [$F_{(2,19)}$ = 0.32, $p$=0.72, $\eta^2$=0.02)] were not significant, but the interaction *Sentence type* X *SNR* was significant [$F_{(4,76)}$ = 13.29, $p$=0.000, $\eta^2$=0.41)]. The triple interaction *Sentence type* X *SNR* X *Age* was not significant [$F_{4,76)}$ = 0.04, $p$=0.998, $\eta^2$=0.00)].

The second level of analysis investigated the degree of conformity of individual patterns with the trends described in the group analysis. For this purpose, the proportion of participants who had score within the total group average ± 1 standard deviation was computed, as shown in Table 1. In general, individual patterns show a certain level of congruence with the total group results, close to the expected proportion of 68 % for a normal distribution of scores. At a SNR of 0 dB, however, the proportion of participants who obtained a correct score within ± 1 standard deviation of the total group average on HP and LP sentences is lower than 60 %, reflecting a greater heterogeneity of the sample. In this case, interpretation of group results may not accurately reflect performance of each individual in the group because of either very low or very high scores.

**Table 1:** Average percentage correct key word recognition for each age group with standard deviation (in parentheses) and difference score at SNRs of -2, 0 and +2 dB, and total group average. For each measure, the proportion of participants with scores within total group average ± 1 standard deviation is indicated.

| | **9-10 years old** | **11-12 years old** | **Total** | **Proportion of score within ± 1 sd from the total group mean** |
|---|---|---|---|---|
| | | | SNR = -2 dB | |
| HP | 33 % (±13) | 40 % (±16) | 37 % (±15) | 13/21 (61 %) |
| LP | 21 % (± 8) | 25 % (±17) | 23 % (±14) | 18/21 (86 %) |
| DS | 12 % (±13) | 16 % (±13) | 14 % (±13) | 13/21 (61 %) |
| | | | SNR = 0 dB | |
| HP | 63 % (±11) | 68 % (±11) | 65 % (±11) | 12/21 (57 %) |
| LP | 40 % (±12) | 41 % (±13) | 40 % (±12) | 11/21 (52 %) |
| DS | 23 % (±12) | 27 % (± 8) | 25 % (±10) | 13/21 (61 %) |
| | | | SNR = +2 dB | |
| HP | 86 % (±10) | 90 % (± 5) | 88 % (± 8) | 15/21(71 %) |
| LP | 56 % (± 8) | 59 % (±15) | 58 % (±12) | 13/21 (61 %) |
| DS | 30 % (± 9) | 32 % (±13) | 31 % (±11) | 13/21 (61 %) |

## Effect of SNR and the linguistic contextual information

Because there was no indication of developmental effects on speech perception in noise between the two age groups (9-10 and 11-12 years), the data were pooled with the 48 remaining participants, for a total sample of 69 children. These 48 additional participants were exposed to three different SNRs (between -6, -4, -2, 0, +2 and +4 dB) to obtain data on a wider range of noise conditions. Results are summarized in Figure 1. A consistent finding is that the mean average of the correct scores for the HP sentences is higher than for the LP sentences at all the SNRs tested, by at least 7 % (at SNR=-3 dB) and up to 28 % (SNR=+2 dB), illustrating the contribution of the linguistic contextual information to auditory speech perception.

Psychometric functions of the group average score (in percentage) for HP and LP sentences as a function of SNR were fitted using a cumulative normal distribution metric with two free parameters (midpoint, slope). The functions illustrated in Figure 2 represent the fitted average score data as a function of SNR for both sentence types as well as the fitted average difference score. The average SNR at which 50 % of the key words were recognized was +0.8 dB for LP sentences and –1 dB for HP sentences. This 1.8 dB difference in SNR to achieve 50 % correct key word recognition in LP and HP sentences represents the contribution of the linguistic contextual information to auditory speech perception for children 9-12 years in age with the TPB.

Another way to look at the contribution of linguistic context in the TPB is by means of the difference score between performance on the HP and LP sentences. The difference in score between the psychometric functions for HP and LP sentences is plotted as a function of SNR in Figure 2. The SNR for which a maximum is reached on this function represents the noise condition where listeners benefit the most from the use of the linguistic contextual cues. In Figure 2, a maximum fitted difference of scores of 20 % is observed at a SNR of 0.8 dB. This group analysis is useful to observe general trends, but it does not necessarily reflect typical performance of individuals in the group. As seen on Figure 1, standard deviations were quite large at some SNRs, probably due to either very low or very high scores.

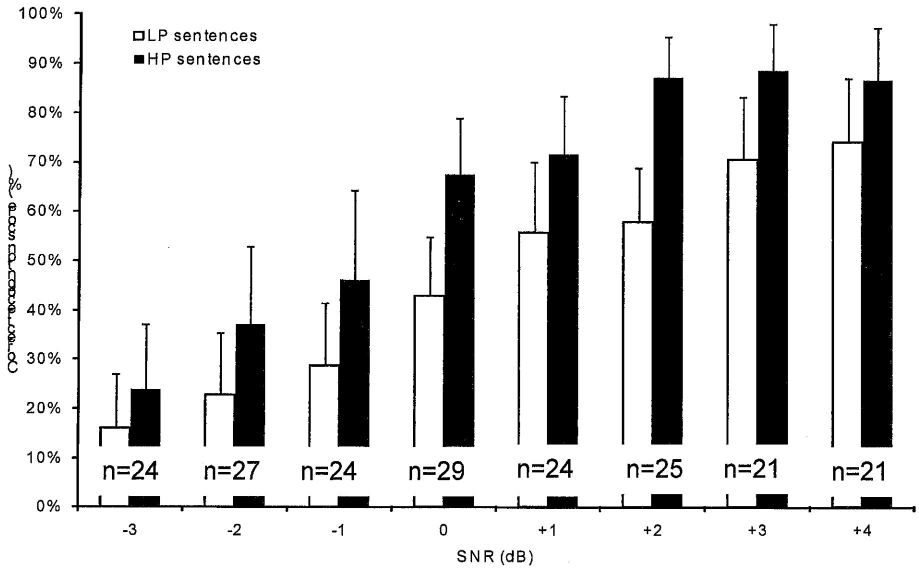


**Figure 1:** Average percentage correct score for LP sentences (white columns) and HP sentences (black columns) at the different SNRs tested, with the sample of 69 children from 9 to 12 years of age.

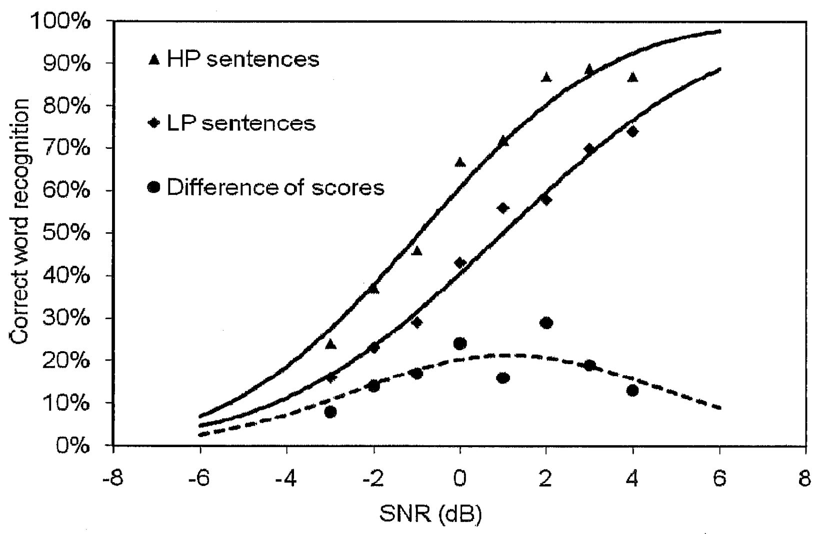


**Figure 2:** Psychometric function of the percentage of key word recognition for HP sentences (triangle symbol), LP sentences (square symbol) and difference score (round symbol) as a function of signal-to-noise ratio (SNR) (N=69).

## Cluster analysis

Cluster analysis on the individual psychometric functions was conducted to capture specific profiles within the children group. To insure uniformity in the computation of the psychometric functions, the analysis was conducted on participants who were exposed to the same number of experimental conditions. As such, the data of 13 participants were removed as they were tested at only two SNR conditions (as opposed to three conditions) because of a lack of time during the experimental phase. Also, to prevent outliers forming clusters with small number of cases, the data from two other participants were taken out prior to the cluster analysis.

The hierarchical cluster analysis was run on the remaining 54 cases using four variables extracted from the psychometric functions for each participant: the value of the SNR threshold to reach 50 % correct score on the LP and HP functions, and the slope of the LP and HP functions. The analysis, conducted using Ward's method, produced two clusters of participants for which some variables were significantly different, as described in Table 2.

The first cluster (n= 25) is characterized by a lower SNR threshold than the second cluster (n= 29) for a 50 % correct score on the LP sentences (-0.6 dB versus 1.9 dB respectively) a difference of 2.5 dB between groups. The first cluster has also a lower SNR threshold for a 50 % correct score on the HP sentences (-1.8 dB) compare to the second cluster (-0.3 dB), a difference of 1.5 dB. The slope of the psychometric functions for the HP and LP sentences were similar.

**Table 2:** Descriptive table of the two clusters that emerged from the data

| | Cluster | N | Mean | Standard deviation | F | Sig.* |
|---|---|---|---|---|---|---|
| Age | 1 | 25 | 133 months | 11 | 2.981 | .090 |
| | 2 | 29 | 128 months | 9 | | |
| | Total | 54 | | | | |
| SNR LP | 1 | 25 | -0.6 dB | 2.00 | 36.84 | .000* |
| | 2 | 29 | 1.9 dB | 0.91 | | |
| | Total | 54 | | | | |
| SNR HP | 1 | 25 | -1.8 dB | 1.01 | 28.275 | .000* |
| | 2 | 29 | -0.3 dB | 0.96 | | |
| | Total | 54 | | | | |
| Slope LP | 1 | 25 | .26 dB/SNR | 0.10 | 0.576 | .451 |
| | 2 | 29 | .28 dB/SNR | 0.13 | | |
| | Total | 54 | | | | |
| Slope HP | 1 | 25 | .37 dB/SNR | 0.12 | 0.098 | .756 |
| | 2 | 29 | .38 dB/SNR | 0.11 | | |
| | Total | 54 | | | | |

A one-way ANOVA was conducted to confirm which classifying variables were statistically different between the groups. There was a significant difference between the group means of the SNR for the LP sentences [$F_{(1,53)}$ = 36.839, $p$=0.000)] and the SNR for HP sentences [$F_{(1,53)}$ = 28.275, $p$=0.000)]. The group mean was not significantly different between the two clusters for the variables age and slope value for the LP sentences and for the HP sentences.

A set of statistical analysis was done to explore the benefit of linguistic context between the two clusters. A one-way ANOVA was conducted on the maximum HP-LP difference score and the SNR value at which the maximum difference score was observed for each case in the two clusters. The results for the maximum difference score did not reach significance level [F = 3.484, $p$ = .068] between the first cluster (mean: 25 %) and the second cluster (mean: 30 %), suggesting that both groups showed comparable gains from linguistic context when listening to speech in background noise. However, a significant difference [F = 7.812, *p* = .007] was observed for the value of the SNR at which the maximum difference score was measured between the two groups (first cluster: SNR of 0.4 dB, second cluster: SNR of 1.6 dB), indicating that participants in the second cluster required a more favourable SNR than those in the first cluster to get comparable benefit from linguistic context.

## CONCLUSION

The results of the present study should be interpreted with caution since only children with normal hearing function without developmental condition participated in the experiment. However, there is an indication that a dynamic approach is needed when analysing the performance of children on speech perception in noise tests. Two subgroups of children, not related to age, emerged from the sample. Both groups showed comparable benefits from use of linguistic context when listening to speech in background noise, but the first cluster required a less favourable SNR than Group 2 to achieve comparable performance.

The SPIN-alike tests, such as the TPB, mimics real life listening conditions. Just as there are some situations in children's life where linguistic and contextual information are available and redundant, there are some situations where there is not as much redundancy, for example, when listening to a new topic in the classroom. As reported by Picard & Bradley (2001), many studies have reported noise levels in excess optimal conditions for understanding speech in the classroom, bringing less than ideal levels of SNRs. For example, Blair (1977) reported SNR values in the classroom ranging from -7 to 0 dB and Finitzo-Hieber (1988) reported SNR values ranging from +1 to + 4 dB. Listening at such SNRs would be very challenging for the group of participants of the present study, especially for the children in the second cluster who would not greatly benefit from the linguistic context because of too low SNR. The situation could be even worse for children with learning difficulties who require higher SNR to understand speech when in presence of noisy background (Bradlow et al. 2003) or bilingual children (Bovo & Callegari 2009) who may not benefit as much from linguistic contextual information.

## ACKNOWLEDGMENTS

This work was supported by a doctoral fellowship from the Centre de recherche du CHU Sainte-Justine and Fonds québécois de la recherche sur la nature et les technologies. The author would also like to thank the Faculty of Health Sciences of the University of Ottawa for a bursary covering the travel expenses for the ICBEN conference.

## REFERENCES

American Speech-Language-Hearing Association (ASHA) (1997). Guidelines for Audiologic Screening [Guidelines]. Available from www.asha.org/policy.

Blair J (1977). Effects of amplification, speechreading, and classroom environment on reception of speech. Volta Review 79: 443-449.

Bloom L, Lahey M (1978). Language development and language disorders. New-York: John Wiley.

Bovo R, Callegari E (2009). Effects of classroom noise on the speech perception of bilingual children learning in their second language: Preliminary results. Audiol Med 7: 226-232.

Bradlow AR, Kraus N, Hayes E (2003). Speaking clearly for children with learning disabilities: Sentence perception in noise. J Speech Lang Hear Res 46: 80–97.

Choi S, Lotto A, Lewis D, Hoover B, Stelmachowicz P (2008). Attentional modulation of word recognition by children in a dual-task paradigm J Speech Lang Hear Res 51(4):1042-54.

Crandell CC, Smaldino JJ (2000). Classroom acoustics for children with normal hearing and with hearing impairment. Lang Speech Hear Serv Sch 31: 362-370.

Elliott LL, Connors S, Kille E, Levin S, Ball K, Katz D (1979). Children's understanding of monosyllabic nouns in quiet and in noise. J Acous Soc Am 66: 12-21.

Fallon M, Trehub SE, Schneider BA (2000). Children's perception of speech in multitalker babble. J Acous Soc Am 108: 3023-3029.

Fallon M, Trehub SE, Schneider BA (2002). Children's use of semantic cues in degraded listening environments. J Acous Soc Am 111: 2242-2249.

Finitzo-Hieber T (1988). Classroom acoustics. In: Roeser R (ed): Auditory disorders in school children, 2nd ed. (pp 221–223). New York: Thieme-Stratton.

Friedrich M, Friederici AD (2004). N400-like semantic incongruity effect in 19-month-olds: processing known words in picture contexts.J Cogn Neurosci 16 (8): 1465-77.

Grant KW, Seitz PF (2000). The recognition of isolated words and words in sentences: Individual variability in the use of sentence context. J Acoust Soc Am 107 (2): 1000-1011.

Kacprzyk JB, Wesam Ashour Wu Y (2009). Non-standard parameter adapatation for exploratory data analysis. Berlin:Springer.

Kalikow DN, Stevens KN, Elliott LL (1977). Development of a test of speech intelligibility in noise using materials with controlled word predictability. J Acoust Soc Am 61: 1337-1351.

Lagacé J, Jutras B, Giguère C, Gagné JP (2011). Speech perception in noise: Exploring the effect of linguistic context in children with and without auditory processing disorder. Int J Audiol 50: 385–395.

Leibold LJ, Neff DL (2007). Effects of masker-spectral variability and masker fringes in children and adults. J Acoust Soc Am 121(6):3666-76.

Mayo LH, Florentine M, Buus S (1997). Age of Second-Language Acquisition and Perception of Speech in Noise. J Speech Lang Hear Res 40: 686-693.

Perrin F, Grimault, N (2005). Fonds sonores. Laboratoire Unités Mixtes de Recherche, Centre National de la Recherche Scientifique 5020, Lyon, France.

Picard MS, Bradley JS (2001). Revisiting speech interference in classrooms. Audiology 40: 221-244.

Rönnberg J, Samuelsson S, Lyxell B (1998). Conceptual constraints in sentence-based lipreading in hearing-impaired. In: Campbell R, Dodd B, Burnham D (eds) : Hearing by Eye II: The Psychology of Speechreading and Audiovisual Speech (pp 143-153). Mahvah: Erlbaum.

Talarico M, Abdilla G, Aliferis M, Balazic I, Giaprakis I, Stefanakis T, Foenander K, Grayden DB, Paolini AG (2007). Effect of age and cognition on childhood speech in noise perception abilities. Audiol Neurootol 12(1):13-9.