



OCENA KLIMATU AKUSTYCZNEGO W KLASACH SZKOLNYCH ZE WZGLĘDU NA ZROZUMIAŁOŚĆ MOWY

Evaluation of the acoustical climate in classrooms with respect to speech intelligibility

Andrzej Ossowski, Julia Smirnowa

Institut Podstawowych Problemów Techniki, PAN
aossow@ippt.gov.pl; julsmir@ippt.gov.pl

STRESZCZENIE

W niniejszej pracy rozpatrzono problem jakości klimatu akustycznego w klasach szkolnych. Zaproponowano funkcję jakości klimatu akustycznego do oceny klas pod względem zrozumiałości mowy.

1. POJĘCIE JAKOŚCI KLIMATU AKUSTYCZNEGO

W pracy [10] przedstawiona została ogólna koncepcja klimatu akustycznego w klasach szkolnych. Klimat akustyczny w klasach z jednym źródłem hałasu, z danym szumem tła oraz źródłem mowy określono za pomocą zestawu wskaźników $\{SA, SU, NA, I_{N0}, I_{S0}, I_B\}$ opisujących własności samej klasy (wzmocnienie mowy SA , czystość mowy SU , wzmocnienie hałasu NA) oraz własności źródeł (intensywność źródła mowy I_S , źródła hałasu I_N oraz szumu tła L_B). Ponieważ klimat akustyczny jest pojęciem wielowymiarowym, nie jest oczywisty sposób jego oceny, który umożliwiłby wybór klasy najlepiej nadającej się do spełnienia określonych wymagań użytkowych (np. zrozumiałości mowy). Porównywanie jakości różnych klas ze względu na określone własności użytkowe wymaga, więc wprowadzenia odpowiedniej, skalarnej funkcji jakości, którą dodatkowo można by uśredniać po miejscach uczniowskich w klasie. Ponieważ zwykle własności użytkowe klasy wyrażane poprzez efekty nauczania (a w szczególności - zrozumiałość mowy) są własnościami bardziej jakościowymi niż ilościowymi, raczej trudno jest podać ich opis ilościowy. Z tego powodu nie znaleziono jak dotąd zadawalających funkcji jakości odpowiadających choćby zrozumiałości mowy w klasach. Funkcja taka powinna istnieć, ponieważ można ją pośrednio szacować na podstawie wyników testów psychologicznych.

Opisana koncepcja klimatu akustycznego (warunek *zupełności* [10]) bazuje w istocie na założeniu, że obiektywne kryterium, które odpowiada określonej własności użytkowej klas powinno być funkcją wskaźników klimatycznych. Jeśli ta hipoteza jest prawdziwa a rozpatrywany zestaw wskaźników - zupełny, to dla każdej własności użytkowej klas powinna istnieć odpowiednia funkcja $K(SA, SU, NA, I_{N0}, I_{S0}, I_B)$, której wartości porządkowałyby klasy wg jakości ze względu na tą własność.

Jest jasne, że funkcja taka może zależeć dodatkowo od pewnych parametrów psychologicznych czy psychoakustycznych. Parametry takie można by jednak niezależnie identyfikować eksperymentalnie. Wtedy taka funkcja jakości byłaby faktycznie wyliczalna tylko na podstawie obiektywnych wskaźników klimatycznych.

2. CZYNNIKI ISTOTNE DLA ZROZUMIAŁOŚCI MOWY

Dysponując pełnym - jak się wydaje - zestawem wskaźników klimatu akustycznego, możemy pokusić się o znalezienie przybliżonej postaci funkcji jakości klimatu akustycznego w odniesieniu do dowolnej cechy użytkowej klas. Poniżej wykonamy to zadanie w odniesieniu do zrozumiałości mowy, która jest stosunkowo dobrze zbadana i dostępnych jest na jej temat wiele konkretnych wyników badań psychoakustycznych [1,2,3,5]. W tym celu musimy najpierw wyszczególnić czynniki warunkujące zrozumiałość mowy. Na podstawie dostępnej wiedzy, możemy wymienić trzy takie czynniki fizyczne i psychoakustyczne, mianowicie: relacja sygnału do szumu (*SNR*), poziom sygnału mowy oraz poziom hałasu w danym miejscu w klasie, przyspieszający efekt zmęczenia uczniów. Wielkości te można wyrazić jako następujące funkcje wskaźników klimatycznych:

poziom użytecznego i nieużytecznego sygnału mowy

$$L_S = 10 \log \frac{I_S}{I_*} = 10 \log \frac{SA \cdot I_{S0} \cdot (r_{S0})^2}{(r_S^{(d)})^2 \cdot I_*},$$

$$L_{SN} = 10 \log \frac{I_{SN}}{I_*} = 10 \log \frac{SA \cdot I_{S0} \cdot (r_{S0})^2}{SU (r_S^{(d)})^2 \cdot I_*} \quad (3)$$

poziom hałasu od źródła punktowego

$$L_N = 10 \log \frac{I_N}{I_*} = 10 \log \frac{NA \cdot I_{S0} \cdot (r_{S0})^2}{(r_S^{(d)})^2 \cdot I_*}, \quad (4)$$

stosunek użytecznego sygnału mowy do szumu

$$SNR = \frac{SU}{1 + \frac{NA \cdot SU}{SA} \cdot \frac{I_{N0} \cdot (r_{N0})^2 \cdot (r_S^{(d)})^2}{I_{S0} \cdot (r_N^{(d)})^2 \cdot (r_{S0}^S)^2} + \frac{SU}{SA} \cdot \frac{I_B}{I_{S0}} \cdot \frac{(r_S^{(d)})^2}{(r_{S0}^S)^2}} \quad (5)$$

gdzie r_S , r_N są odległościami miejsca percepcji dźwięków odpowiednio od źródła mowy i źródła hałasu, natomiast r_{S0} , r_{N0} są promieniami tych sfer ograniczających odpowiednie źródła, na których powierzchni określone są intensywności I_{S0} , I_{N0} [10].

Funkcja jakości klimatu akustycznego w klasach ze względu na zrozumiałość mowy powinna być więc zależna od powyższych wskaźników. Naszym zadaniem będzie określenie postaci tej zależności.

3. JAKOŚĆ KLIMATU AKUSTYCZNEGO ZE WZGLĘDU NA ZROZUMIAŁOŚĆ MOWY

Jeśli przyjąć, że wymienione wyżej czynniki wpływają na zrozumiałość mowy niezależnie (tzn. nie występują między nimi synergistyczne lub antagonistyczne zależności), to poszukiwany wskaźnik jakości klimatu akustycznego klas powinien mieć postać iloczynu trzech funkcji zależnych odpowiednio od: stosunku sygnału do

szumu, bezwzględniego poziomu sygnału mowy oraz poziomu hałasu, przy czym do hałasu należy zaliczyć nie tylko sygnał pochodzący od źródła ale również szum tła oraz nieużyteczny sygnał mowy.

Nie zmniejszając ogólności możemy przyjąć, że rozpatrywany wskaźnik jakości powinien przyjmować wartości w przedziale $[0,1]$ przy czym wartość 1 (i odpowiednio 0) będzie oznaczać maksymalną (minimalną) zrozumiałość mowy. W szczególności, wartość 1 (lub 0) funkcji zależnej od danego czynnika będzie oznaczała, że przyczyny go określające nie mają negatywnego wpływu na zrozumiałość mowy (lub uniemożliwiają jakąkolwiek zrozumiałość). Poniżej określone zostaną konkretne postaci poszczególnych funkcji składowych wydedukowane na podstawie dostępnych danych doświadczalnych.

Jest oczywiste dla $SNR = 0$ niemożliwa jest percepcja sygnału mowy a więc i zrozumiałość mowy. Praktycznie istnieje pewna minimalna wartość wskaźnika $SNR_0 > 0$, poniżej której niemożliwa jest percepcja mowy i to niezależnie od poziomu użytecznego sygnału mowy. Dla $SNR \approx SNR_0$ tylko najsilniejsze formanty w słowach są poprawnie percypowane. Ze wzrostem wartości SNR , gdy kolejne fonemy stają się prawidłowo percypowanymi a określone grupy słów danego języka - zrozumiałymi, zrozumiałość mowy powinna rosnąć, przy czym dla dużych wartości SNR ten przyrost powinien być coraz mniejszy. Ponieważ teoretycznie wartość SNR może być dowolnie duża, trzeba przyjąć, że zrozumiałość mowy rośnie asymptotycznie do pewnej wartości maksymalnej, gdy $SNR \rightarrow +\infty$. Rozpatrywana funkcja powinna więc asymptotycznie dążyć do jedności, co oznacza, że w zakresie dużych wartości SNR wskaźnik ten nie wywiera istotnego wpływu na zrozumiałość mowy. Funkcją spełniającą opisane wyżej wymagania jakościowe może być np. funkcja wykładnicza postaci

$$E = 1 - \exp[-(SNR - SNR_0)/\alpha] \quad (6)$$

(określona dla $SNR \geq SNR_0$), w której SNR wyraża się wzorem (5). Wartości SNR_0 i parametru $\alpha > 0$ można oszacować za pomocą niezależnych eksperymentów psychoakustycznych lub na podstawie danych literaturowych [2,7]. Wielkość parametru α determinuje bowiem granicę, powyżej której można uznać wartości wskaźnika SNR za wysokie. Powinna to być taka wartość SNR , dla której wszystkie fonemy danego języka są słyszalne i mogą być prawidłowo rozpoznawane, jeżeli nie ma ku temu innych przeszkód. Ponieważ mechanizm maskowania mowy szumem i odwrotnie zależy od widmowych gęstości mowy i szumu, wszystkie wartości ze wzoru (6) trzeba szacować w paśmie mowy. Składowe hałasu z poza tego pasma nie wpływają bowiem na proces percepcji mowy. Zakłada się, że hałas w klasie szkolnej ma widmo zbliżone do szumu szerokopasmowego. Gdyby ten hałas zawierał wyraźne piki np. tony, wielo-tony czy impulsy, wtedy oszacowanie składowych SNR powinno być przeprowadzone w pasmach np. w pasmach krytycznych czy *ERB* ucha [9].

Drugim zagadnieniem wymagającym omówienia jest zależność zrozumiałości mowy od poziomu sygnału mowy. Wiadomo, że percepcja sygnału mowy niezależnie od poziomu zakłóceń (a więc również niezależnie od wielkości wskaźnika SNR), jest możliwa tylko w pewnym zakresie poziomu sygnału użytecznego. Na pewno sygnały poniżej progu słyszalności i powyżej granicy bólu nie są percypowane a tym bardziej - rozumiane [9]. Zbyt wysokie poziomy sygnału powodują, bowiem przesterowania komórek nerwowych odpowiedzialnych za percepcję, natomiast poziomy zbyt niskie nie są w stanie pobudzić tych komórek. Praktycznie uzasadnione jest przyjęcie istnienia pewnych granicznych poziomów użytecznego sygnału mowy - poziomu dolnego, określonego przez natężenie I_1 oraz poziomu górnego, określonego przez I_2 .

Czynnik w poszukiwanym wskaźniku jakości, określający zależność zrozumiałości mowy od poziomu sygnału powinien więc maleć do zera w pobliżu granicznych poziomów sygnału i tym samym - osiągać maksimum (równe 1) wewnątrz przedziału $[I_1, I_2]$. Łatwo sprawdzić, że najprostszą funkcją spełniającą powyższe warunki jest następująca funkcja

$$F = -4 \cdot \log[SA \cdot (I/I_1)] \cdot \log[SA \cdot (I/I_2)] / [\log(I_2/I_1)]^2 \quad (7)$$

którą umownie nazwiemy *rozpoznawalnością mowy*. Oznaczając przez I_0 standardowe natężenie odniesienia, możemy powyższą funkcję wyrazić wprost za pomocą poziomu mowy $L_S = 10\log(I/I_0)$ oraz poziomów krytycznych L_1, L_2 , odpowiadających intensywnościom I_1, I_2 . Łatwo sprawdzić, że funkcja F jest następującą, kwadratową funkcją poziomu użytecznego sygnału mowy

$$F = -4[L_S - L_1 + \log(SA)] \cdot [L_S - L_2 + \log(SA)] / [L_2 - L_1]^2 \quad (8)$$

Z danych literaturowych wynika, że dla dużych poziomów sygnału (>69dB) zrozumiałość mowy maleje ze wzrostem poziomu sygnału natomiast dla mniejszych poziomów (w granice 50-55 dB) – rośnie [13]. Oznacza to, że optymalny (ze względu na zrozumiałość mowy) poziom mowy jest rzędu 60dB lub nieco mniejszy. Być może nie przypadkiem taki właśnie poziom uznaje się jako typowy dla mowy nauczyciela w klasie [2,6,8]. Jeśli przyjąć, że maksymalny poziom L_2 jest poniżej progu bólu a więc np. 110 dB a poziom optymalny równy 60 dB, wówczas łatwo wyliczyć na podstawie (6), że poziom minimalny L_1 powinien być równy 10 dB, a więc nieco powyżej progu słyszalności. Uzyskana zgodność oszacowania wartości krytycznych i wartości optymalnej poziomu mowy oraz zgodność jakościowa z danymi doświadczalnymi uzasadnia wybór funkcji postaci (6), (7). Dokładniejsze wartości poziomów krytycznych L_1, L_2 można dobrać uzgadniając przebieg funkcji (7) z dostępnymi danymi dla poziomów mniejszych i większych od optymalnego.

Ostatnią kwestią do omówienia jest sprawa zmęczenia towarzyszącego procesowi percepcji mowy. W trakcie wykładu układ słuchowy człowieka oraz wyższe piętra układu nerwowego stopniowo obniżają swoją zdolność percepcji i rozumienia mowy. Dzieje się tak nawet w przypadku mowy czystej bez jakiegokolwiek hałasu czy pogłosu. Obecność zakłóceń może tylko przyspieszać proces zmęczenia [4,6,14]. Zrozumiałość mowy jest więc na ogół funkcją czasu a rozważany czynnik powinien asymptotycznie maleć od wartości maksymalnej równej jedności do zera (co odpowiada zupełnej dekoncentracji ucznia po dłuższym okresie słuchania lekcji). Eksperymenty psychologiczne wskazują, że procesy zmęczenia przebiegają jakościowo podobnie u różnych osób [4,7]. Można się spodziewać, że początkowo, w pewnym okresie zmęczenie ucznia będzie niewielkie a więc percepcja i rozumienie mowy - bliskie maksimum. Dopiero po upływie pewnego czasu nastąpi wyraźny spadek zdolności rozumienia mowy. Oznacza to, że odpowiednia funkcja opisująca zmęczenie powinna mieć punkt przegięcia i zerową pochodną dla czasu $t = 0$. Najprostszą i najdogodniejszą funkcją czyniącą zadość powyższymi wymaganiami jakościowym jest funkcja Gaussa postaci:

$$G = G(t) = \exp[-(t^2/2\tau^2)] \quad (9)$$

gdzie parametr τ (jako dyspersja funkcji Gaussa) określa czas, po którym następuje gwałtowny spadek wydolności ucznia, jeżeli chodzi o zrozumienie mowy. Wartość tą można wyznaczyć z niezależnych eksperymentów psychologicznych. Wiadomo, że w zależności od wieku słuchacza i jego samopoczucia czas ten mieści się zwykle w przedziale od kilkunastu do kilkudziesięciu minut. Głównym składnikiem

parametru τ będzie oczywiście stała czasowa τ_0 określająca tempo procesu zmęczenia ucznia w warunkach percepcji mowy czystej (tj. bez zakłócenia hałasem i pogłosem). Wpływ wymienionych zakłóceń na zdolność rozumienia mowy nie jest zbyt duży i dobrze zbadany. Dlatego, aby uwzględnić zakłócenia w rozpatrywanej funkcji Gaussa wystarczy wprowadzić do naturalnej stałej czasowej τ_0 tylko liniowe poprawki zależne od poziomu hałasu oraz poziomu nieużytecznego sygnału mowy. Zatem stała czasowa zmęczenia ucznia będzie wyrażać się wzorem

zmodyfikowana stała czasowa procesu zmęczenia:

$$\tau \approx \frac{\tau_0}{1 + \beta \cdot L_N + \gamma \cdot L_{SN}}, \quad (10)$$

gdzie parametry β i γ określają odpowiednio negatywny wpływ hałasu L_N i nieużytecznego sygnału mowy L_{SN} na proces zmęczenia ucznia. Równość parametrów β , γ oznacza jednakowe traktowanie wszystkich zakłóceń. Można się spodziewać, że współczynnik γ okaże się większy od β , ponieważ nieużyteczny sygnał mowy bardziej niż hałas utrudnia rozumienie sygnału użytecznego.

Zbierając razem powyższe formuły możemy, zaproponować następującą funkcję jakości klimatu akustycznego klas szkolnych ze względu na zrozumiałość mowy:

Funkcja jakości klimatu akustycznego

$$K = E \cdot F \cdot G = E(SNR) \cdot F(L_S, SA) \cdot \exp\left[-(t^2/2\tau^2)\right]. \quad (11)$$

Jak widać, jakość klimatu akustycznego w klasach szkolnych ze względu na zrozumiałość mowy jest dość złożoną funkcją, zależną od wszystkich wskaźników klimatycznych SA , NA , SU oraz od poziomów sygnału mowy, hałasu i tła. Złożona postać funkcji K , całkowicie odmienna od dotychczas rozpatrywanych w literaturze, praktycznie wyklucza jej przypadkowe odgadnięcie. Funkcja ta może posłużyć do optymalizacji adaptacji akustycznej i przybliżonej oceny klas szkolnych ze względu na zrozumiałość mowy. Można też oczekiwać lepszej korelacji wartości funkcji K z wynikami testów psychologicznych na zrozumiałość mowy.

Warto zauważyć, że wskaźnik jakości klimatu jest iloczynem trzech czynników E , F , G , z których dwa pierwsze mogą być bliskie jedności w klasach o dużych wartościach wskaźnika SNR oraz poziomach użytecznego sygnału mowy bliskich poziomowi optymalnemu (50-60 dB). W takich przypadkach początkowa zrozumiałość mowy będzie praktycznie jednakowa tj. słabo zależna od wartości wskaźników klimatycznych klasy. Może to tłumaczyć problemy skorelowania wyników psychoakustycznych testów słownych z tymi wskaźnikami. Czas trwania testów jest bowiem zwykle zbyt krótki na to, aby wystąpiły wyraźne objawy zmęczenia słuchaczy. Czynnikiem zmęczenia może więc odgrywać kluczową rolę w optymalizacji i ocenie akustyki typowych klas szkolnych. Przyjmując czas trwania wykładu jako T , powinniśmy stosować do oceny jakości klasy średnią wartość funkcji jakości w przedziale $\langle 0, T \rangle$. Oznacza to, że właściwym wskaźnikiem oceny jakości klas jest funkcja $J = E \cdot F \cdot \Phi(t/\tau)$, gdzie $\Phi(t/\tau)$ jest dystrybuantą standaryzowanego rozkładu normalnego.

LITERATURA

1. J. S. Bradley, Speech intelligibility studies in classrooms, *J. Acoust. Soc. Am.*, 80, 846-854, (1986).
2. J. S. Bradley, H. Sato, Speech Intelligibility Test Results for Grades 1, 3 and 6 Children in Real Classrooms, in proceedings of ICA, Kyoto, (2004).
3. S. R. Bistafa., J. S. Bradley, Reverberation time and maximum background-noise level for classrooms from a comparative study of speech intelligibility metrics, *J. Acoust. Soc. Am.*, 107, 861-875 (2000).
4. G. W. Evans, S. J. Lepore, Nonauditory effects of noise on children: A critical review. *Children's Environments*, 10(1), pp. 31-51, (1993).
5. M. Hodgson, Rating, ranking, and understanding, acoustical quality in university classrooms, *J. Acoust. Soc. Am.*, 112, 568-575, (2002).
6. Z. Koszarny, P. Goryński, Narażenie uczniów i nauczycieli na hałas w szkole, *Rocz. PZH*, 5 – 6, 257 – 310, (1990).
7. P. H. Lindley, D. A. Norman, *Procesy przetwarzania informacji u człowieka – wprowadzenie do psychologii*, PWN, Warszawa, (1984).
8. D. J. MacKenzie, S. Airey, *Classroom Acoustics: A Research Project*, Heriot-Watt University, Edinburgh, (1999).
9. B. J. Moore, *Podstawy psychologii słyszenia*, PWN Warszawa – Poznań, (1999).
10. A. Ossowski, J. Smirnowa, Koncepcja klimatu akustycznego w klasach szkolnych, w materiałach konferencyjnych *Otwarte Seminarium z Akustyki*, Gdańsk, (2004).
11. S. Sakamoto, Y. Suzuki, S. Amano, T. Kondo, Speech intelligibility by use New word-lists with controlled word familiarities and a phonetic balance, in proceeding of ICSV, Hong Kong, (2001).
12. B. R. Slater, Effect of noise on pupil performance. *J. Education. Psychol.*, 59(40), 239-243, (1968).
13. G. A. Studebaker, R. L. Sherbecoe, Monosyllabic word recognition at higher-than-normal speech and noise levels, *J. Acoust. Soc. Am.* 105(4), 2431 – 2444, (1999).
14. L.C. Sutherland, D. Lubman, The impact of classroom acoustics on scholastic achievement, in proceeding of the 17th International Congress on Acoustics, Rome, (2001).