

Évaluation automatique de la parole spontanée en anglais langue étrangère : le rôle des pauses et de l'accent lexical dans la compréhensibilité du locuteur

Sylvain Coulange

Résumé long (4000 caractères max. espaces compris)

La compétence de production orale en langue étrangère est communément évaluée à travers la capacité du locuteur à se faire comprendre. Le terme de compréhensibilité est souvent utilisé dans le domaine de l'acquisition des langues pour désigner le degré d'effort requis par un auditeur pour comprendre le locuteur. Celle-ci dépend de nombreux facteurs, côté locuteur comme auditeur, parmi lesquels la position des marqueurs d'hésitation et le rythme de la parole sont souvent décrits comme facteurs clés dans les grilles d'évaluation. En revanche, les approches automatiques reposent encore majoritairement sur des comparaisons à un modèle, souvent à partir de parole lue, sans porter une attention particulière aux phénomènes linguistiques susceptibles de perturber la compréhension.

Dans cette thèse, nous proposons un outil d'évaluation automatique de la production orale spontanée en anglais, ciblant spécifiquement deux phénomènes linguistiques susceptibles d'impacter la compréhensibilité du locuteur : la distribution syntaxique des pauses et l'accentuation lexicale. Cet outil a été utilisé pour analyser la variation de ces phénomènes chez des locuteurs de niveau B1 et B2 francophones et japonophones (L2), ainsi que chez des locuteurs anglophones natifs (L1). Nous avons ensuite mesuré leur impact sur la perception de l'effort de compréhension en temps réel chez des auditeurs anglophones natifs.

Les analyses révèlent qu'une différence majeure entre les locuteurs B1 et B2 d'une part, et entre les L1 et L2 d'autre part, réside dans la fréquence des pauses de bas niveau syntaxique : plus le niveau de compétence augmente, plus les pauses tendent à se concentrer aux frontières syntaxiques de haut niveau (par exemple, entre les propositions). À partir de ces observations, nous avons défini un score de distribution syntaxique des pauses pour caractériser cette concentration.

Concernant l'accentuation lexicale, les mesures font ressortir une influence marquée des patterns accentuels de la langue maternelle des locuteurs. Les francophones augmentent généralement la hauteur (f_0) et allongent la dernière syllabe des mots, tandis que l'intensité reste stable sur l'ensemble des syllabes. Bien que les locuteurs B1 et B2 se chevauchent largement, une progression significative est observée entre ces deux niveaux : les locuteurs B2 accentuent plus fréquemment la syllabe attendue et produisent un contraste acoustique plus marqué en termes de f_0 et d'intensité. En revanche, les locuteurs japonais, dont la langue maternelle intègre un accent lexical, positionnent l'accent avec davantage de précision et produisent des contrastes prosodiques plus importants que les francophones.

Enfin, pour mesurer l'impact des différents types de pauses et de patterns d'accentuation lexicale sur la perception de la compréhensibilité, nous avons recruté 60 auditeurs afin d'évaluer dynamiquement 16 extraits de paroles issus du corpus des locuteurs francophones mentionné ci-dessus. Lors de l'écoute, il leur était demandé de cliquer sur un bouton dès qu'ils ressentaient un effort pour comprendre le locuteur, quelle qu'en soit la raison. L'analyse des patterns de clics, une fois normalisés, indique une augmentation significative de la difficulté perçue entre 0 et 3 secondes après une pause située à l'intérieur d'un syntagme, ainsi qu'entre 2 et 3 secondes après un mot dont la syllabe accentuée n'était pas celle attendue. À l'inverse, la difficulté perçue diminue significativement après une pause située entre deux propositions ou après un mot correctement accentué. Ces observations renforcent l'idée que les pauses de bas niveau syntaxique et les mots accentués de manière inattendue impactent directement la perception de difficulté de compréhension. Notre expérience démontre que cet impact est direct et mesurable.

ENGLISH VERSION

Automated Assessment of Spontaneous Speech in English as a Foreign Language: the Role of Pauses and Lexical Stress in Speaker Comprehensibility

Oral production proficiency in a foreign language is commonly assessed through the speaker's ability to be understood. The term comprehensibility is frequently used in the field of second language acquisition to refer to the level of effort required by a listener to understand the speaker. Comprehensibility depends on numerous factors, both speaker- and listener-related, among which the position of hesitation markers and speech rhythm are often highlighted as key factors in evaluation scales. In contrast, automatic assessment approaches still predominantly rely on comparisons with a model, often based on read speech, without specific attention to linguistic phenomena likely to disrupt comprehension.

In this thesis, we propose a tool for the automatic evaluation of spontaneous oral production in English, specifically targeting two linguistic phenomena likely to impact speaker comprehensibility: the syntactic distribution of pauses and lexical stress. This tool was used to analyze variation of these phenomena among French- and Japanese-speaking learners of English at B1 and B2 proficiency levels (L2), as well as among native English speakers (L1). We then measured their impact on the perception of real-time comprehension effort in native English-speaking listeners.

The analyses reveal that a major difference between B1 and B2 speakers, as well as between L1 and L2 speakers, lies in the frequency of low-syntactic-level pauses: as proficiency increases, pauses tend to concentrate more at high-syntactic-boundary locations (e.g., between clauses). Based on these observations, a syntactic pause distribution score was developed to quantify this concentration.

Regarding lexical stress, the findings show a strong influence of speakers' native language stress patterns. French speakers typically increase pitch (f_0) and lengthen the final syllable of words, while intensity remains stable across syllables. Although B1 and B2 speakers largely overlap, significant improvement is observed between these levels: B2 speakers more frequently stress the expected syllable and produce stronger acoustic contrasts in terms of f_0 and intensity. On the other hand, Japanese speakers, whose native language incorporates lexical stress, position stress more accurately and produce greater prosodic contrasts than French speakers.

Finally, to measure the impact of different types of pauses and lexical stress patterns on perceived comprehensibility, we recruited 60 listeners to dynamically evaluate 16 speech excerpts from the aforementioned corpus of French speakers. While listening, participants were asked to click a button whenever they experienced difficulty understanding the speaker, for whatever reason. An analysis of the normalized click patterns indicates a significant increase in perceived difficulty between 0 and 3 seconds following a pause within a phrase, and between 2 and 3 seconds after a word with an unexpected stressed syllable. Conversely, perceived difficulty significantly decreased after a pause between clauses or following a correctly stressed word. These findings support the idea that low-syntactic-level pauses and unexpected lexical stress directly impact the perception of comprehensibility. Our experiment shows that this impact is direct and measurable.

Résumé grand public (1000 caractères max. espaces compris)

Comprendre un locuteur qui s'exprime dans une langue étrangère demande parfois un effort, notamment lorsque le rythme ou les pauses dans la parole paraissent peu naturels. Cette thèse explore les phénomènes linguistiques qui influencent la compréhension, c'est-à-dire l'effort requis pour comprendre un locuteur. En analysant des conversations spontanées en anglais produites par des francophones et des japonophones, un outil a été développé pour étudier la position des pauses et l'accentuation des mots. Les résultats montrent que les locuteurs plus expérimentés placent leurs pauses à des endroits plus appropriés, comme entre les propositions, et accentuent mieux les syllabes attendues. Une analyse de l'effort perçu en temps réel révèle que ces caractéristiques réduisent l'effort de compréhension par les auditeurs anglophones natifs. Cette recherche ouvre la voie à des outils d'apprentissage qui aident à améliorer la compréhension des locuteurs en langue étrangère.

ENGLISH VERSION

Understanding a speaker talking in a foreign language sometimes requires effort, especially when the rhythm or pauses in their speech seem unnatural. This thesis explores the linguistic phenomena that influence *comprehensibility*, defined as the effort required to understand a speaker. By analyzing spontaneous English conversations produced by French and Japanese speakers, a tool was developed to investigate pause placement and word stress. The results show that more experienced speakers place pauses more appropriately, such as between clauses, and produce more accurate lexical stress. An analysis of perceived effort in real time reveals that these features reduce comprehension effort by native English-speaking listeners. This research opens the way to learning tools that help improve the comprehensibility of foreign language speakers.