

Title	日本語音声の振幅包絡に含まれる冗長性に関する検討
Author(s)	西野, 恭生
Citation	
Issue Date	2014-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/12054
Rights	
Description	Supervisor: 鷓木祐史, 情報科学研究科, 修士

日本語音声の振幅包絡に含まれる冗長性に関する検討

西野 恭生 (1110047)

北陸先端科学技術大学院大学 情報科学研究科

2014年2月12日

キーワード: 日本語音声, 振幅包絡, 冗長性, 変調成分.

音声は知覚的に頑健である。音声の言語情報（なんとしゃべっているのか）の取得に関する重要な特徴は、音声の至るところに分散して冗長に含まれている。そのため、その一部が変形・削除されたとしても我々は容易に言語情報を取得することができる。その一例として、雑音駆動音声の知覚が知られている。雑音駆動音声とは、音声の振幅包絡を取り出して、その振幅包絡で雑音を振幅変調することにより、時間微細構造を持たず、音声の振幅包絡情報のみを持った信号である。この雑音駆動音声の先行研究より、音声の了解性に関する情報は、音声の時間包絡線（振幅包絡）に多く含まれていることが報告されている。しかし、これらの先行研究では、振幅包絡に冗長性があるのかどうかまでは深く踏み込んで検討されていない。そのため、音声の振幅包絡に含まれる冗長性を解明する。

音声の冗長性とは、音声知覚に関係した特徴が音声信号に分散して冗長に含まれており、その中のいくつかが失われたとしても何と言っているのか正確に聴き取ることができる性質である。この性質を利用した技術として、音声圧縮や音声符号化技術がある。これらは、ヒトの錯覚を利用し聴き取りに影響を与えない特徴を、オリジナルの音声信号から削除・変形させている。そこで、振幅包絡の特徴を削除や変形といった制御をしても言語情報の取得に影響のない特徴に着目しなければならない。その特徴として、変調成分に着目した。振幅包絡の変調成分の周波数が $4 \sim 16$ Hz は音声了解度に重要であり、変調成分の周波数が $2 \sim 8$ Hz に特有のピークを持つと報告されている。しかしながら、この特徴を変化させたときに言語情報の取得にどのような影響を与えるのかは検討していない。そのため、振幅包絡に含まれる変調成分を系統的に制御し、言語情報の取得が可能であるかどうか調査する。

振幅包絡に含まれる変調成分を系統的に制御するにあたって、音声のモーラの時間構造とモーラの時間構造よりも短い時間長の子音の時間構造に着目した。モーラの時間構造に着目した聴取実験では、1モーラの時間長を1周期としたときに得られる変調成分の上限周波数が言語情報の取得が可能である境界であることが分かった。子音の時間構造に着目した聴取実験では、子音を表現する振幅包絡を再現するだけの変調成分が必要と考え、変調成分の上限周波数は高いと予測した。しかし、聴き取り可能だった変調成分の上限周

波数は低かった。これは、子音部の振幅包絡を再現しなくとも言語情報の取得が可能であることを示している。子音の言語情報の取得が可能であったのは、子音部の振幅包絡の立ち上がりにより言語情報の取得を可能としていたと考えられる。

言語情報の取得に必要な変調成分は、モーラの時間構造の観点からは、モーラの時間構造を再現できるだけの変調成分が保存されていればよく、子音（摩擦音）の観点からは、振幅包絡の立ち上がりの違いによって、言語情報の取得が可能であることが分かった。

これらの結果から、低域通過フィルタにより削除された変調成分が日本語音声の振幅包絡に含まれる冗長性だと解釈できる。