12-2018

Segmental Contributions to Speech Intelligibility in Nonconcatenative vs. Concatenative Languages

Yahya Aldholmi
University of Wisconsin-Milwaukee

SEGMENTAL CONTRIBUTIONS TO SPEECH INTELLIGIBILITY IN NONCONCATENATIVE VS. CONCATENATIVE LANGUAGES

by

Yahya Aldholmi

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Linguistics

at

The University of Wisconsin-Milwaukee

December 2018

ABSTRACT

SEGMENTAL CONTRIBUTIONS TO SPEECH INTELLIGIBILITY
IN NONCONCATENATIVE VS. CONCATENATIVE LANGUAGES

by

Yahya Aldholmi

The University of Wisconsin-Milwaukee, 2018
Under the Supervision of Professor Anne Pycha

This study investigated the contributions of segments (consonants vs. vowels) to speech intelligibility in Arabic and English. In these two languages, consonants and vowels play crucially different grammatical roles. Arabic is a nonconcatenative language that primarily assigns lexical information to consonants and morphosyntactic information to vowels, while English is a concatenative language that does not assign distinct roles to either class of segments. On this basis, we hypothesized that consonants and vowels would play very different roles in the intelligibility of the two languages. Five laboratory experiments were conducted, three on Arabic and two on English. Participants listened to words and sentences in which either all consonants or all vowels were replaced with silence and were asked to indicate what they heard. Unlike previous studies, all stimuli were carefully controlled for ratio of consonants to vowels. Results showed that in Arabic, consonants made a greater contribution than vowels to speech intelligibility, both in isolated words and in complete sentences. Furthermore, in consonant-only conditions, stimuli containing more consonants were

more intelligible than those containing more vowels, displaying a clear effect of segmental ratio. In English, by contrast, the consonants and vowels made roughly equivalent contributions to speech intelligibility, and segmental ratio played a negligible role. These two disparate findings suggest that segmental contributions are crucially modulated by language-specific factors. That is, the different contributions of consonants and vowels to speech intelligibility are not solely determined by their distinct acoustic cues, but also by the grammatical role they play.

DEDICATION

إهــداء

To my dear mother who kept waiting for ten years to receive a joyous answer to her persistent question: "When will you be back?" And to my dear father who had that very question but kept hiding it for the same length of time.

To my stoical wife who assisted with undertaking all of my concerns while she also was burdened with her own.

To my two little daughters who lived an expatriate life by no choice of their own.

To my late brother whom I lost in my childhood and I will miss indefinitely.

To my brothers and sisters who thrived while I was away.

To my instructors who believed in my capabilities.

To all my friends who supported me.

إلى والدتي العزيزة التي مكثتْ عشر سنينَ ترقُب إجابة سارّة لسؤالها الدائم: متى ستعود إلينا؟ ووالدي العزيز الذي لبثَ عشرًا يُضمر السؤال نفسه.

إلى زوجتي الصبورة التي شاطرتني كل همومي بينما كانت هي الأخرى مثقلة بهمومها الخاصة.

إلى طفلتيّ الصغيرتين اللتين عاشتا إلى جانبي غربةً لا ناقة لهما فيها ولا جمل.

إلى أخي الراحل الذي فقدته في طفولتي وأفتقده منذئذٍ إلى يومئذٍ.

إلى أشقائي وشقيقاتي الذين ترعرعوا في غيابي.

إلى أساتذتي الذين آمنوا بقدراتي.

إلى أصدقائي الذين دعموني.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

2C-1V	2 Consonants to 1 Vowel
2C-2V	2 Consonants to 2 Vowels
2C-3V	2 Consonants to 3 Vowels
3C-1V	3 Consonants to 1 Vowel
3C-2V	3 Consonants to 2 Vowels
3C-3V	3 Consonants to 3 Vowels
4C-1V	4 Consonants to 1 Vowel
4C-4V	4 Consonants to 4 Vowels
ASIN	Arcsine
CI	Confidence Interval
CO	Consonant-Only
CR	Correct Response
CV-Ratio	Consonant-to-Vowel Ratio (in the inventory of a given language)
df	Degree of Freedom
Exp(B)	Exponentiation of the *B* coefficient (Odds Ratio)
F1, 2, 3	Formant 1, 2, and 3
FI	Final
GEE	Generalized Estimating Equations
IN	Initial
M	Mean
MED	Median
MLE	Maximum-Likelihood Estimation
ms.	Millisecond
MSA	Modern Standard Arabic
N	Noun

Q-Q Quantile-Quantile

QIC Quasi Likelihood under Independence Model Criterion

QICC Corrected Quasi Likelihood under Independence Model Criterion

RAU Rationalized Arcsine Unit

S-Ratio Segmental Ratio (in speech)

SE Standard Error

STD Standard Deviation

TIMIT Texas Instruments/Massachusetts Institute of Technology Database

V Verb

VO Vowel-Only

LIST OF SYMBOLS

:	Ratio (When Placed between Two Values)
/%/	Percentage/Proportion
/ð/	Voiced Dental Fricative
/ə/	Schwa
/ɛ/	Open-Mid Front Unrounded Vowel
/ɣ/	Voiced Velar Fricative
/ħ/	Voiceless Pharyngeal Fricative
/ɪ/	Near-High Front Unrounded Vowel
/ŋ/	Velar Nasal
/ɹ/	Alveolar Approximant
/ʃ/	Voiceless Postalveolar Fricative
/x/	Voiceless Velar Fricative
/ʒ/	Voiced Postalveolar Fricative
/ʔ/	Glottal Stop
/ʕ/	Voiced Pharyngeal Fricative
/ˤ/	Emphaticness/Pharyngealization
/θ/	Voiceless Dental Fricative
≈	Approximately
x	Interaction of Different Factors
χ^2	Chi-Square
ω^2	Omega-Squared (for Effect Size)

ACKNOWLEDGEMENTS

A dissertation is officially attributed to one author but shepherded by multiple individuals, and this dissertation is no exception. I truly owe a debt of gratitude to more people than I could possibly name in a few pages. I have encountered and addressed several dilemmas in this dissertation, but I am unsure how to approach this final dilemma. I will do my best to list and thank as many people as I can for their unforgettable support. If anyone has been left out, I offer my profuse apologies.

First and foremost, I would like to express my heartfelt thanks to my dissertation supervisor and committee chair, Prof. Anne Pycha, without whose guidance this work would have never come to fruition. Pycha, as a linguistics professor, and I, as a MA linguistics student, both arrived at the Department of Linguistics at UW-Milwaukee (UWM) in 2012. The first class I enrolled in with her was a required course titled *Introduction to Sounds and Sound Systems*. In that class and on the first day, I introduced myself to the class and she quickly posed a question to me: *What brought you to UWM?* My response was: *Well, a combination of reasons, but mainly because the Department has a faculty member who is interested in Arabic syntax*. I admit now, six years later, that was not quite an astute answer. My initial impression of Dr. Pycha, which I stress still holds true, was: *This is a genius (and probably difficult) professor*. Later, I opted to enroll in multiple courses with her during my MA and PhD programs and learned from her more than one could ever infer from this dissertation alone.

I gradually became more and more eager to continue learning from her, so I approached her about chairing my preliminary examination which she welcomed warmly. During my preliminary work I experienced some challenges, especially because I was pursuing (and overwhelmed with) a Master's degree in computer science and information technology, about which Prof. Pycha did not know at that time. However, she was without a shadow of doubt successful in helping me eliminate blocks and obstacles and navigate the intricacies of the preliminary examination climate, thus progressing towards my dissertation phase. Her unwavering support and guidance during my preliminary examination, both of which continued later on, encouraged me to ask her to serve as my dissertation chair. Prof. Pycha, a seasoned researcher, transformed me from an inexperienced and unprepared student into an independent and confident researcher. Throughout my dissertation, Prof. Pycha was able to balance wholehearted support and constructive critique. Every time she noticed from afar that I was drowning in the workload and gasping for a breath of air, she pulled me up to the surface. Prof. Pycha was promoted to associate professor and took a well-deserved yearlong sabbatical leave before I completed this dissertation. Notwithstanding, I never felt any difference in her availability or responsiveness during that time; she continued to guide my work and provide instant feedback until I dotted the last "i" in this dissertation. Hence, for this reason and many others not cited here, I am so grateful to Prof. Pycha. Note to future students: Prof. Pycha is a science writer and professional editor, so be prepared to sharpen your writing skills!

I am also thankful to the committee members of this dissertation (listed alphabetically): Prof. Fred Eckman, Hamid Ouali, and Hanyong Park. Each of the three was an equally potential advisor for my dissertation should it have been on a different topic. If there existed a legitimate way to have each of them chair a chapter of the dissertation, you would have seen them all listed as my supervisors. Prof. Eckman was like a loving and caring father not only to me but to all students, especially to foreign students who always thought of his office as a second home. In the retirement celebration that the Department held to honor Prof. Eckman, not only did current students attend the event to listen to him talk one more time, but also several alumni came from different states and even from different countries. I will leave it to you, the reader, to imagine all the tears that welled up in the audience's eyes and the whirlwind of emotions that erupted during and after his talk.

Prof. Ouali was a splendid supervisor for my MA thesis in syntax, and I am extremely honored to have had him as a contributing co-author in one of my publications. Within a few days after I set foot in Milwaukee in 2012, Prof. Ouali, who was the Department Chair, invited me to visit him in his office to put a face to the name I had been corresponding with before my arrival. He was the first faculty member who welcomed me to the Department. Prof. Oauli introduced me to the Arabic Linguistics Society and Symposium, encouraged me to participate, and helped me to stand on my feet and grow up professionally. Likewise, Prof. Park was one of the few early people who pushed me to participate in linguistics conferences and helped me develop my research skills. He introduced and accompanied me to several conferences

such as the MidPhon and the American Acoustics Society conferences. I am honored to have worked with him on several projects and to have a prolific co-author of his caliber in one of my publications. Prof. Park was not formally a co-supervisor for my dissertation, but he always offered his assistance in the resolution of phonetics, acoustics, and statistics problems. He also provided detailed feedback on the dissertation after the defense took place, which made the work better. I am truly proud to have all three of these figures as my additional committee members.

The entire department's faculty at UWM is phenomenal. They are all true exemplars for the type of instructors, advisors, and researchers I had always hoped to someday meet, work with, and learn from. Therefore, I am also thankful to all other faculty members and academic staff in the department (listed alphabetically): Prof. Edith A Moravcsik, Garry Davis, Jae Yung Song, Nicolas Fleisher, Sandra Pucci, and Tue Trinh, as well as the administrative staff in the department: Sharon Geibel and (previously) Teri Mannes. I am also thankful to other faculty members who taught me in other courses during my program, namely: Mrs. Annie Mayers (Educational Statistics Department), Holly Smith (English Department), Prof. Razia Azen (Educational Statistics Department), Sara Held (Mount Mary University), Prof. Sheryl Slocum (Alverno College), Prof. Wendi Halstead (Milwaukee Area Technical College), and Yael Gal-Ben Yitschak (Department of Foreign Languages and Literature - Hebrew). I am also grateful to many other instructors who taught me a variety of other subjects such as mathematics, statistics, computer science, information technology, instructional design, educational technology, Middle Eastern studies, and Hebrew at

both UWM and other schools. They always understood my grueling schedule as a student pursuing multiple degrees at the same time. It would be my pleasure to list all of those great people here, but I hope that they will accept my sincere apologies for not being able to do so in such a limited space.

I must also thank my Milwaukeean network of classmates, colleagues, and friends who made my foreign student's life much smoother and by far less stressful: Abdel-Rahman Abu Helal, Abdellatif Oulhaj, Amal El Haimeur, Amara Sankhagowit, Bara Omari, Christine Mary, Daniel Phillips, Dola Algady, Dylan Pearson, Ho Eun Park, Islam Hindi, Jacklyn Ryan, Juman Al-Bukhari, Khuloud Labnieh, Laura Kiskunas, Laura Johanna Mendoza-Montanez, Maria Liya Mar, Maria Teresa Bonfatti, Mario Sanchez Gumiel, Nicole Desjarlais, Robin Fritche, Robyn Sperling, Siriporn Lerdpaisalwong, Vivian Watson Jiang, Wafaa Khaleel (my daughter's Arabic language private teacher), and Zafer Al-Lababidi. I am particularly thankful to the Saudi community members who created a home-like atmosphere and native-feeling environment for me and my little family in Milwaukee: Abdullah Alsubhi, Abdulrahman Aljutaili, Ali Alzayid, Bader Alharbi (who received his PhD before I did, so Dr. Bader Alharbi now), Hamsa Badra, Khalid Alfaifi, Khalid Dandani, Malak Bakhshwain, Mohammed Abuhaib (my obliging neighbor in Shorewood), Saad Alamri, Saad Alshahrani, Salem Albuhayri, Salman Albardi, Turki Alwahibee, and Yasser Albaty, as well as their families. They all motivated me to work hard and complete my degree and in particular Dr. Bader Alharbi and Salem Albuhayri kept me optimistic with their ardent conversations about my progress every so often.

Those who were present on my defense day (November 27th, 2018), supported me, and celebrated the event also deserve special thanks: Abdel-Rahman Abu Helal, Abdullah Alsubhi, Dong Jin Kim, Jieun Lee, Dr. Kelsie Pattillo, Mohammed Abuhaib, Prof. Nicolas Fleisher, Robin Fritche, Salem Albuhayri, Salman Albardi, and Yasser Al-baty. On December 9th, 2018, a few days before I left the United States, my Saudi friends held an auspicious celebration and farewell event, and overwhelmed me with their sublime gifts and delicate feelings to the extent that thank-yous were almost all I could say during the entire event. Thank you again to all of you, and particularly to Abdulrahman Aljutaili who organized the event.

Hundreds of other friends from all over the world and especially from Saudi Arabia, the United Kingdom, and the United States have been witnessing my educational journey and foreseeing its completion. They are all thanked, but it would necessitate that I stay in Milwaukee for an extended time beyond graduation to come up with a full list of their names. I have to name two close and special friends from Saudi Arabia, who supported me from the time I was an undergraduate student there all the way to now, and have been waiting for the moment I inform them about my graduation. Those are Ahmad Alkhusaafi, an Arabic language instructor, and Ali Aldholmi, an English language instructor. I cannot thank them enough, but they are both aware of how much I appreciate their tremendous assistance and unremitting support. Two other friends whom I always shared concerns with and received support from are Essa Alfaifi, a PhD candidate at the University of Arizona, and Dr. Yahya Aldholmi, a professor at Al-Imam Mohammed University. My friends, Mohammed

Al-Salmani and Yahya Algobaisy, were always of great help while I was a student in the United Kingdom and the United States. Other linguistics and non-linguistics friends whose names are now floating to mind are all thanked, but I apologize for not mentioning them name by name. It is heartening that they will understand and let me turn my attention to my family members.

I am indebted to my dear parents, Jaber Aldholmi and Jameelah Alkhusaafi. My father had a dream to continue his education more than a half century ago, but he was limited by endless constraints that even most capable people of his generation did not escape. Every time he narrated his life circumstances to me, he unknowingly nourished that very dream in me until his dream became mine. I am grateful to him for the inspiration, motivation, and strength he instilled in me. My mother, on the other hand, was the main individual who insisted nonstop that a better education would guarantee a fuller and brighter life and thus encouraged me to pursue my undergraduate studies. In 2008, I left my home country in pursuit of graduate education, and since then she has been praying for me and waiting for the second I convey good news to her. Thank you so much for all you did for me. In fact, I am unsure that I would have reached this level of education if it were not for the two of you. My thanks also extend to my brothers and sisters who have always been proud of me and have been hoping for the day I complete this degree. I especially thank my brother, Ahmad Aldholmi, for taking care of my parents over the past ten years while I was away.

My own little family, Asma Alfaifi, Bailasan Aldholmi, and Ihsaas Aldholmi, also contributed to my success. What I will articulate here about my wife, Asma Alfaifi, is not because a husband and wife should compliment one another, but because in this case a husband and a wife have complemented each other. Asma stood by my side day after day, hour after hour during my long journey of education although she, too, was a busy student and had her own assignments, exams, projects, papers, and conference presentations. Asma was also able to teach me many aspects of her subject (Information Science & Information Technology) during her undergraduate and graduate studies, although I was not successful enough at teaching her linguistics. I am eternally grateful for her sheer tenacity, stubborn persistence, personal sacrifices, incredible courage, and infinite patience. Without her, I would have struggled mightily and would not have achieved the success that I have had so far, including the completion of this dissertation

Bailasan, my smart and adult-like daughter, has always cared about her parents' progress, even with the little life experience she has as a kindergarten-aged child. Her primary goal in Milwaukee was to help me and her mother to finish our studies and return to our home country. Whenever she noticed me involved in a task (reading, typing, or thinking), she avoided interrupting me. Sometimes, she wanted to play games with me, but rather than asking to do so directly, she would instead politely ask: *Dad, do you have homework today?* If the answer was yes, which in fact was the case most of the time, she would accept it and leave calmly. In the last two years of my PhD and when we had a new child, Ihsaas, it always impressed me how

Bailasan attempted to keep her little sister happy so that she did not interrupt me. Bailasan had a critical question: *Are we going back to Saudi Arabia forever this time?* Her mother, Asma, had another critical one: *Are you pursuing another degree?* I am so glad that my answers to the two questions have now become "yes" and "no", respectively, after they used to be "no" and "yes" for years. I also thank my little source of laughter and tears, Ihsaas. I cannot deny that because of her our schedules have gotten busier and messier, but she has also brought a new joy to the house. Huge thanks to the three of you!

Finally, thank you to all those who supported me financially. King Saud University, represented by the Saudi Arabian Cultural Mission (SACM) in the United States, generously funded my education for ten years (2008-2018). SACM itself funded my wife's studies and helped her accomplish two degrees, something that my wife and I truly appreciate. The Department of Linguistics also funded my travel to give talks and present research papers at different conferences. I am grateful for their support, and I particularly thank Prof. Gary Davis, the department chair during that period of time. My thanks extend to the graduate school for their financial support, as well.

⇐
Table of
Contents

CHAPTER ONE:

INTRODUCTION

⇒
Chapter
Two

As the two major classes of segments in human speech, both consonants and vowels contribute information to the architecture of words and sentences (i.e., speech). These two types of segments are phonetically very different from one another. While consonants are realized with some degree of stricture in the vocal tract, vowels are not. Understanding the role of each segment type in lexical formation and how much information is conveyed by each type in words and sentences can aid in modeling *speech intelligibility*, defined as the degree to which a listener comprehends a speaker's message[1] (see e.g., Henningsson, Kuehn, Sell, Sweeney, Trost-Cardamone, & Whitehill, 2008) or, more simply, how well a listener understands speech. This topic (i.e., the contribution of vowels vs. consonants) has been explored in a small set of languages, including English (Cole, Yan, Mak, Fanty, & Bailey, 1996), Spanish, Dutch (Cutler, Sebastián-Gallés, Soler-Vilageliu, & van Ooijen, 2000), and Mandarin Chinese (Chen, Wong, & Wong, 2013; Chen, Wong, Zhu, & Wong, 2015). In general, the results of these studies have shown that consonants contribute more than vowels to speech intelligibility at the word level (except in Chinese), whereas vowels contribute more than consonants to speech intelligibility at the sentence level (including

[1] Note that similar definitions have been used in the literature interchangeably to refer to closely related concepts such as understandability and acceptability. The above-stated definition will be used when referring to speech intelligibility throughout this dissertation.

Chinese). For instance, a Dutch word such as *komeet* "come" is more likely to be recognized in a sentence when only its vowels *–o–ee–* are made available to listeners, as compared to a condition in which only its consonants *k–m–t* are presented. The reverse is true, however, when the word is presented in isolation.

The majority of languages in which this topic has been investigated are concatenative systems in which phonological forms have a relatively even distribution of consonants and vowels, and morphological forms are created by combining discrete stems with discrete affixes. However, languages can differ greatly in the way they employ and distribute consonants and vowels in their morphological systems. This presents major challenges to the development of a universal and cross-linguistic theory about speech intelligibility. While languages such as English, Spanish, and Dutch do not assign fundamentally different roles to vowels and consonants in the morphology, other languages do. For example, in an English word such as "akin" /əkɪn/, the two consonants /k, n/ are morphologically equivalent to the two vowels /ə, ɪ/ in terms of how the word is built, as English does not functionally distinguish between consonants and vowels when constructing its morphemic units. In nonconcatenative languages such as Arabic, on the other hand, we typically see a key morphological distinction between consonants and vowels, as will be elucidated below.

1.1 Segmental Differentiability in Nonconcatenative Languages

Semitic languages such as Arabic and Hebrew use a nonconcatenative morphological system (also referred to as non-linear or root-pattern morphology) in which a sequence of consonants signals lexical (semantic) information, whereas vowels are

intercalated to signal structural (morphosyntactic) information, similar to the function of affixes in concatenative languages (McCarthy, 1979). An example paradigm is provided in Table 1, where we see that the Modern Standard Arabic (MSA) consonantal root ʕ-ʃ-q "related to LOVE"[2] can be used to generate different lexical items by interdigitating this consonantal root with a string of vowels (often referred to as a pattern). The consonantal skeleton typically consists of 3 consonants (although a sequence of two or four consonants is also possible) and forms a discontinuous and unpronounceable morpheme on one tier to supply the core or general semantic information around which all related forms are clustered. A separate pattern composed mainly of vowels acts as another morpheme on a separate tier to supply morphosyntactic information such as ontological categories for nouns and aspectual properties for verbs (see Holes, 2004; McCarthy, 1981; Watson, 2007, among many others[3]). Note that patterns, which usually consist of vowels, are not wholly represented in Arabic orthography except for some teaching and learning purposes; this is also the case in Hebrew (Feldman, Frost, & Pnini, 1995).

[2] Capitalization here means that the English word represents the abstract general meaning of the consonants independent of any morphosyntactic information.
[3] Other studies have attempted to propose a concatenative morphology in Semitic languages, for example Benmamoun 2003 and Ratcliffe (1998), but I do not subscribe to such views (yet).

Table 1. Example of root-pattern interdigitating in Modern Standard Arabic

Lexical Item	Meaning	Form
/ʕ**a**ʃ**i**q/	"Loved"	Perfective/past
/ʕ**aa**ʃ**i**q/	"Lover"	Agent/doer/active participle
/ʕ**a**ʃ**ii**q/	"Beloved"	Adjective
/ʕ**a**ʃ**uu**q/	"Who is in deep love"	Adjective
/ʕ**u**ʃʃ**aa**q/	"Who are in deep love"	Plural of active participle
/ʕ**a**ʃ**a**q/	"Love/Loving"	Base form/gerund/nominal

The meaning of a Semitic word is collectively constructed via the root and pattern combination, sometimes in an idiosyncratic fashion (Bentin & Feldman, 1990; Feldman, Frost, & Pnini, 1995). As the paradigm above illustrates, the semantic information associated with LOVE is shared across the different forms, but each has a different structure reflected by a change in the quantity and/or quality of vowels. The number of consonants and vowels in the various forms is generally fixed and predictable. For example, a basic (not causative or reciprocal) active perfective form will have 3 consonants and 2 vowels, as in /ʕaʃiq/ "loved", and the consonant-to-vowel ratio will remain 3:2 for all other roots conjugated in the active perfective, as in /karih/ "hated", /fariħ/ "got happy", /ħazin/ "got sad", /saxitˤ/ "got furious", /nadim/ "regretted", and so on. A passive form of these examples will change only the quality but not the quantity of the first vowel, as in /ʕuʃiq/ "was loved" and /kurih/ "was hated", whereas the agentive form will feature a change in the quantity but not the quality of the first vowel (3C-3V ratio) as in /ʕaaʃiq/ "lover" and /kaarih/ "hater". An Arabic speaker would not

predict a triconsonantal root such as /ʕ-ʃ-q/ to have a basic perfective form that has a 3C-3V ratio or any regular agentive form to have a 3C-2V ratio. This consonant-to-vowel ratio predictability is special to nonconcatenative languages and is not observed in concatenative languages.

1.2 Motivation

In surveying the literature, the closest previous studies have come to examining the effect of segment's ratio in speech intelligibility was to compare languages that contained different numbers of consonants vs. vowels in their phonemic inventory (henceforth, language CV-Ratio) as in the case of Spanish (5 vowels vs. 20 consonants) vs. Dutch (16 vowels vs. 19 consonants) (e.g., Cutler et al., 2000). Although these studies suggest that the ratio of segments across languages plays no role, the relative morphological contributions of segments within a language has not been investigated yet. In addition, segmental ratio (henceforth, S-Ratio), which is meant to refer to the ratio of consonants to vowels in a given speech sample (either a word or a sentence), has not been investigated at all. For instance, it seems plausible to think that a word like "thrift" with a 4C-1V ratio would be responded to very differently than a word like "any" with a 1C-2V ratio, yet to our knowledge, no studies of speech intelligibility have manipulated this variable.

In sum, if consonants in nonconcatenative languages carry lexical information while vowels carry structural information (see e.g., Holes, 2004; McCarthy, 1981; Watson, 2007), we would expect that consonantal and vocalic segments each play a distinct role in speech intelligibility, and furthermore that S-Ratio would exert

different effects in nonconcatenative languages vs. concatenative languages.

1.3 Questions and Hypotheses

Our starting point is the observation that a segment's contribution to speech intelligibility may not be solely based on its phonetic properties, but also on its morphological properties, motivating a comparison between concatenative and nonconcatenative languages. Within each language, we also conducted a comparison between words and sentences, because the literature to date reports very different results for these two types of speech input. In addition, the number of vowels and consonants (i.e., S-Ratio) in words and sentences, which was not examined as a potential factor in previous studies, was critical in this current study to test our predictions about segmental contributions. Below, I outline each of these three key predictor factors: nonconcatenative vs. concatenative, words vs. sentences, and balanced vs. imbalanced ratios, in greater detail.

1.3.1 Nonconcatenative vs. Concatenative

The primary investigation of segmental contributions to speech intelligibility in this study involves a comparison between nonconcatenative languages, represented in this study by Arabic, and concatenative languages, represented by English.

There is ample evidence to show that speakers of nonconcatenative languages use consonantal roots to recognize, access, and retrieve Arabic words (e.g., Abu-Rabia, 1997; Bentin & Feldman, 1990; Deutsch & Frost, 2003; Deutsch, Frost, Pollatsek, & Rayner, 2000; Feldman, 2000; Holes, 2004; Minouni, Kehayia, Jerema, 1998; Ravid & Shimron, 2003). For instance, priming studies have shown that words that share

the same consonantal root such as /ʕ**a**ʃ**i**q/ "loved" and /ʕ**a**ʃ**ii**q/ "beloved" prime each other; furthermore, they prime each other more than words which share only semantic relatedness such as /ʕ**a**ʃ**ii**q/ "beloved" and /ħabiib/ "beloved". This apparent primacy of consonants leads to the prediction that speech intelligibility benefits more from consonantal information than vocalic information in nonconcatenative languages. Note that similar findings are consistently reported for English – the root (such as "write" in "rewrite" or "writer") is the constituent that English listeners recognize first (see e.g., Caramazza, 1988) and hence will use to comprehend speech in general. The substantial difference between roots in Arabic and in English is that an Arabic root is a consonantal morpheme while English roots can be composed of both vowels and consonants. Together, this raises the question as to whether segmental contributions to speech intelligibility will differ between nonconcatenative and concatenative languages. Therefore, I posed and sought to answer the following question:

- *What are the segmental (vocalic vs. consonantal) contributions to speech (words vs. sentences) intelligibility in nonconcatenative languages, specifically in Arabic?*

I argued that previous findings which emphasized the role of vowels in the intelligibility of sentences and of consonants in the intelligibility of words were crucially restricted to non-Semitic languages. When we turn to Arabic and other nonconcatenative languages, I predict we will see a very different pattern. Specifically, I hypothesize that unlike in concatenative languages, consonants in Arabic are a primary lexical information carrier and thus serve as the fundamental contributing segment to

speech intelligibility in both words and sentences. However, because vowels are responsible for morphosyntactic information in Arabic morphology, vocalic information in sentences carries more structural information that helps listeners to identify the word class (e.g., nominal vs. verbal and passive vs. active) and the word's role within the sentence. This leads to the second issue this study investigated: differences in speech intelligibility with respect to words vs. sentences. This is a crucially important factor, because if Arabic consonants are special, we may expect to see their role modulated differently in words compared to sentences.

1.3.2 Words vs. Sentences

Many of the discrepancies in the previous findings about the roles of vowels and consonants in speech intelligibility are largely a result of the stimulus design choices. Studies such as Cole et al. (1996) used complete English sentences and found that vowels play a larger role than consonants do in speech intelligibility. Others such as Owren and Cardillo (2006) used isolated English words as their stimuli and found the opposite result. More recent studies (e.g., Fogerty et al., 2012) have compared the contribution of vocalic and consonantal segments in both isolated words and complete sentences, and found a greater contribution to speech intelligibility of vocalic information in sentences, but not in words. Considering the way Arabic employs vowels as carriers of structural information, we can predict that vowels in Arabic should provide listeners with a mechanism to recognize some structural information about words in complete sentences and ultimately make speech more intelligible. Nevertheless, such structural information remains a secondary contributor compared to the

lexical information conveyed by consonants. This comparison between words and sentences is also driven by the primary comparison between nonconcatenative and concatenative languages, because unlike in English, structural information in Arabic is tied to a pattern consisting mainly of vowels. Hence, this comparison is still meant to answer the above research question as the notion of speech intelligibility encompasses both words and sentences. Related to this is the reasonable question as to whether S-Ratio can also play a role in speech intelligibility, because if consonants play a primary role in speech intelligibility in Arabic, then their contribution may be modulated more by S-Ratios than is the contribution of vowels. This is the third factor to be addressed by this study, for both languages.

1.3.3 Segmental Ratio: Balanced vs. Imbalanced

One important predictor that tests the importance of consonantal vs. vocalic information is the ratio of consonantal vs. vocalic information in the stimuli presented to the listener. The second question being asked here is as follows:

- *How does segmental ratio (balanced vs. imbalanced) affect speech (words vs. sentences) intelligibility in nonconcatenative vs. concatenative languages (Arabic vs. English)?*

S-Ratio is expected to affect speech intelligibility in Arabic and English, but in different manners. Specifically, the effect of S-Ratio in Arabic is associated with the Arabic morphological system, which is expected to play a crucial role in speech intelligibility as outlined above. Therefore, whenever the ratio limits or neglects consonants, we predict that Arabic speech intelligibility will be affected negatively. This is

not the case in English. In this language, S-Ratio is likely to affect the contributions of both classes of segments to speech intelligibility similarly, simply because English morphemes are composed of consonantal and vocalic sources of phonetic information that play roughly equivalent morphological roles.

As mentioned above, previous studies have almost entirely neglected S-ratio in the construction of their stimuli. When we examine the actual number of vowels vs. consonants used in the stimuli for these studies, S-Ratio seems to have had some effects on their results, as will be detailed in Chapter 2. Consequently, a comparison between Arabic and English with regard to S-Ratio seems both significant and necessary, and is a novel contribution of this dissertation.

These three factors (concatenative vs. nonconcatenative, word vs. sentence, and balanced vs. imbalanced ratio) lead to the three key comparisons of this dissertation. An additional comparison occurs within the experimental technique of noise replacement: in all five experiments presented here, stimulus words or sentences were presented in two forms: vowel-only (henceforth, VO) vs. consonant-only (henceforth, CO).

1.3.4 Vowel-Only vs. Consonant-Only

To facilitate comparison with previous work, we follow a technique commonly used in previous studies, namely replacement with silence. The idea is that, for any given speech input (either an isolated word or a full sentence), two different stimuli are created: one stimulus in which all of the consonants are replaced with silence but the vowels remain (VO condition), and another in which all of the vowels are replaced with silence but the consonants remain (CO condition). For example, for the English

word “banana” /bənænə/, the VO stimulus would be /—ə—æ—ə/, while the CO stimulus would be /b—n—n—/. The participant’s task is to listen to the stimuli and indicate what they hear. Comparing the results from the two types of stimuli allows us to quantify the relative contribution of consonants and vowels to speech intelligibility. This comparison was included in the experiments conducted for this study.

Five experiments were conducted to examine the three factors outlined above. The experiments were organized based on the language they were designed to investigate (hence, Experiments 1, 2, and 3 for Arabic, and Experiments 4 and 5 for English). Below is a full outline of the dissertation.

1.5 Structure

The remainder of the dissertation is organized as follows:

Chapter Two is a literature review in which I summarize the previous findings on the contributions of consonants and vowels to speech intelligibility in concatenative systems, as well as previous studies on speech recognition in nonconcatenative systems.

Chapter Three describes three experiments on Arabic. The first experiment investigated segmental contributions and the contributions of S-Ratios to speech intelligibility at the word level. The second experiment and the third experiment examined segmental contributions and the role of S-Ratio in speech intelligibility both at the sentence level.

Chapter Four describes two experiments on English examining the effect

of S-Ratios on speech intelligibility at the word and sentence levels in English.

Chapter Five is the final chapter in which I provide a general discussion of the entire work, including experimental findings and their overall implications, and conclude the dissertation.

⇐ Chapter One

CHAPTER TWO:

LITERATURE REVIEW

⇒ Chapter Three

As mentioned in Chapter 1, different segments, specifically vowels vs. consonants, carry different types of phonetic information, a fact that has attracted researchers to examine how each type contributes to speech intelligibility. Such research has been conducted in languages such as English and Spanish, but not yet in Semitic languages. However, research on a related but slightly different topic – namely, word recognition – has examined Semitic languages, especially Hebrew, and has shown consistent findings on the importance of the consonantal root in mental representation, lexical access, and word recognition. Hence, in this chapter I review 1) the major previous findings on consonantal vs. vocalic information in speech intelligibility in non-Semitic languages, and 2) the main previous studies on word recognition in Semitic languages.

2.1 Segmental Contributions in Concatenative Languages

Unlike other linguistic constituents such as syllables and phonemes, the difference between vowels and consonants in lexical access and speech recognition has been largely overlooked (van Ooijen, 1996). Theoretical models and approaches to word recognition such as the Cohort Model (Marslen-Wilson, 1978), Trace (McClelland & Elman, 1986), and Shortlist (Norris, 1994) treated both vowels and consonants the same, although some researchers have examined biases for consonants vs. vowels in early word recognition and lexicon development, as in Havy, Serres, and Nazzi

(2014). Likewise, and as pointed out by van Ooijen (1996), experimental studies using single segment manipulation took consonants as their focal point. They mainly used consonant manipulation, and assumed that findings for one category of segments should apply to the other category, too. For instance, Marslen-Wilson & Zwitserlood (1989) examined the importance of word onsets in lexical access, using Dutch and nonsense priming items that differed in the initial consonant such as *honing* "honey" vs. *woning* "dwelling" vs. *foning* (nonsense word). No mention was made of vowels as potential word onsets.

More recent studies, by contrast, have discussed the distinction between vowels and consonants, although most of these were based on Indo-European languages such as English, Spanish, and Dutch. Such studies sparked a debate about which segment type contributes more to speech intelligibility, and presented discrepant findings supporting vowel privilege in some cases and consonant privilege in others. Below, I summarize the major studies based on their outcomes: some studies concluded that vowels contribute more ("vowel's privilege"), some concluded that consonants contribute more ("consonant's privilege"), and yet others concluded that vowels and consonants each contribute more in different situations ("vowel vs. consonant's privilege").

2.1.1 Vowel's Privilege

Cole et al. (1996) examined the relative role of vowels vs. consonants in speech intelligibility. Their study has been influential, and it inspired many subsequent works; therefore, I will explain it in some detail. The researchers classified English

segments into three types: consonants (20 segments), vowels (20 segments), and weak sonorants (12 segments: liquids, glides, and nasals). They conducted three experiments, using 60 sentences spoken by 30 males and 30 females and taken from the Texas Instruments/Massachusetts Institute of Technology- TIMIT Corpus (Garofolo, Lamel, Fisher, Fiscus, & Pallett, 993; Zue, Seneff, & Glass,1990). In Cole et al.'s (1996) first experiment, there was one condition in which vowels (but not weak sonorants) were replaced with either white noise or periodic sounds, and another condition in which consonants (but not weak sonorants) were replaced. There was also one condition without replacement, making up a total of five conditions: no replacement, consonant replaced with white noise, vowel replaced with white noise, consonant replaced with periodic sounds, and vowel replaced with periodic sounds. The participants were asked to listen to each sentence up to five times at their own pace and to write down as many words as they understood. Results showed a significant effect of segment type, although this effect differed for words vs. sentences. Results showed that the type of segment had a significant effect on speech intelligibility, whereas the type of replacement (i.e., white noise vs. periodic sounds) exerted no significant effect. Speech intelligibility was higher when the participants were presented with VO stimuli than when they were presented with CO stimuli.

However, Cole and colleagues still had a concern about the classification of segments. Specifically, the presence of weak sonorants in both the VO and CO conditions was questionable, because it was not clear how much they contributed to the sentence intelligibility in each condition. It was possible that vowels benefited more

from the presence of weak sonorants than consonants did, perhaps because of the coarticulatory information contained in vowels that transition to and from weakened sonorants. Therefore, the researchers conducted a second experiment in which they preserved one class of segments and replaced two classes in each condition. Results ruled out the weak-sonorant explanation. Sentence intelligibility declined dramatically across all three conditions, vowels-only, consonants-only, and weak-sonorants-only.

Another caveat is that co-articulatory information may have been responsible for the better intelligibility in the vowel conditions. It is known that, although vowels and consonants both exert co-articulatory effects on each other, vowels may carry more information about the adjacent consonants due to formant transitions in both the onset and offset of each vowel (see e.g., Cooper, Delattre, Liberman, Borst, & Gerstman, 1952). As a result, vowels carrying more than just their own spectral information may potentially provide more information about speech compared to consonants carrying only their spectral information. However, Cole et al. (1996) investigated this possibility in an additional experiment in which the boundaries of vowels and consonants were either expanded or shrunk by 10 ms from both sides (10*2 = 20 ms) as depicted in Figure 1 below (borrowed from Cole et al. [1996, p. 3] with some design modifications). It was expected that, if vowels carried some local information that informed the speakers about the missing consonants, their performance would decline when vowels were shrunk and increase when consonants were expanded. The

results showed that the reduction or expansion of the time window modulated intelligibility in the consonant condition, but not in the vowel condition.

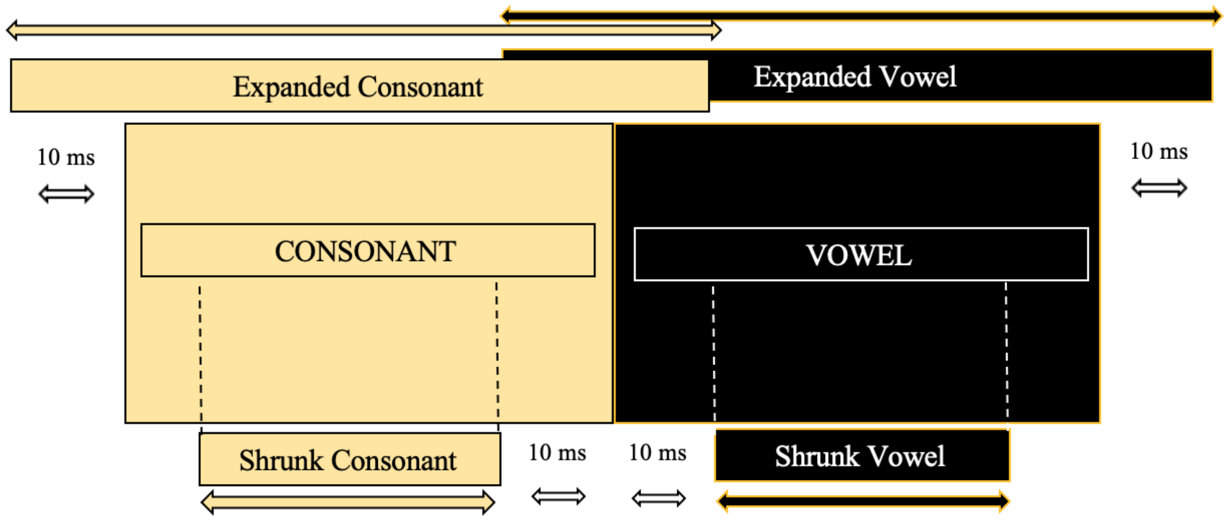


Figure 1. Illustration of expansion vs. reduction of consonant vs. vowel boundaries

On the basis of these three experiments, Cole and colleagues concluded that vowels' contribution to speech intelligibility is greater than that of consonants. However, there are several methodological issues here. First, the classification of segments into 3 classes allowed for an equal number of vowels (20) and consonants (20) and for a comparison between two maximally dissimilar categories by excluding the weak sonorants. Although this exclusion of weak sonorants from the consonants category provided us with two clearly distinct categories, stops and fricatives (consonants) vs. vowels, it is unusual and unjustified. If this classification were based on levels of sonority (consonants vs. weak sonorants vs. vowels), then it is unclear why they did not further divide stops and fricatives into two different categories. Second, the authors reported the number of occurrences of vowels (756), consonants (747),

and weak sonorants (417) in the entire stimuli, but did not provide this information for each individual stimulus sentence. They only claimed, somewhat vaguely, that they obtained almost an equal number of consonants and vowels in each sentence. In fact, however, consonants and vowels were not really balanced; if we add the weak sonorants to the consonant category, the consonants would outnumber the vowels. This imbalance of vowels and consonants in the stimuli seems to be an important issue. The intelligibility of a word that has many more consonants than vowels, such as the English word "strength" /stɹɛŋθ/, would obviously decrease in the VO condition compared to in the CO condition, simply because a large proportion of information is unavailable. Third, although Cole and colleagues did not find a difference between the two types of noise replacement, noise replacement per se was probably not the best technique for distinguishing the contributions of different segment types. White noise can sound like a fricative, for example, potentially leading participants to "re-store" a phoneme that is not present in the signal and thereby affecting perception of the word or sentence. It could also distract the participants, especially when the noise-replaced segments in a word outnumbered the unreplaced segments. All these aspects were fully considered in the current experiments.

Three other studies (Burkle, 2004; Burkle, Kewley-Port, Humes, & Lee 2004; Kewley-Port, Burkle, & Lee, 2007)[4] are in agreement with Cole et al. In Kewley-Port et al. (2007) the performance of normal hearing listeners on sentence intelligibility

[4] The first study is an unpublished Master's thesis undertaken at Indiana University-Bloomington, while the second seems to be a presentation at the Acoustic Society of American conference. I focus here on the published paper by Kewley-Port, Burkle, & Lee (2007), although all three of the studies are listed in the reference section.

was compared to that of hearing-impaired elderly listeners. The noise replacement paradigm used in Cole et al. (1996) was adopted with a small modification, being that the weak sonorants were classified as consonants rather than a separate category. The first experiment replicated the results of Cole et al. (1996), and showed a 40% increase in sentence intelligibility in the VO condition (74%) compared to the CO condition (34%). In the second experiment, although young normal hearing listeners performed better than the hearing-impaired listeners, both groups statistically performed better when only vowels were present in the stimuli, compared to when only consonants were present.

In addition to these studies on English and other Indo-European languages, two recent studies investigated segmental contributions in Mandarin Chinese (Chen et al., 2013; Chen et al. 2015). Mandarin has tones that function phonemically, just as consonants and vowels do, but it does not show a special treatment of vowels vs. consonants. Chen et al. (2013) used sentences from the Mandarin speech perception (MSP) corpus (Fu, Zhu, & Wang, 2011) and adopted the noise replacement technique from Cole et al. (1996), while Chen et al. (2015) used isolated monosyllabic words from a large database. The findings from both studies confirmed the advantage of vowels over consonants in speech intelligibility, which supports the idea that the phonetic characteristics of vowels vs. consonants determine their contribution to speech intelligibility in a universal manner. However, as we will see, not all studies support Cole et al.'s findings regarding the superiority of vowels in English speech intelligibility, and it is this topic that I turn to next. Having introduced and discussed the

major studies supporting vowel's privilege, I now discuss studies which support consonant's privilege.

2.1.2 Consonant's Privilege

Unlike Cole et al., several studies have argued that consonants contribute to speech intelligibility more than vowels do, although the experimental tasks used by these authors often differed from the noise replacement technique. van Ooijen (1996)[5] showed that consonants are more useful than vowels in word recognition for English listeners. Her study used the "reconstruction task", in which participants were given auditory nonsense inputs such as "kebra" and were instructed to rapidly substitute one segment to convert it into a meaningful English word. In the free-choice condition, either the consonant or the vowel could be changed, so that *kebra* can become "zebra" or "cobra". In two other conditions, the subjects could only modify a consonant (in one condition) or a vowel (in the other condition). The results showed more errors in vowel change compared to consonant change. The author concluded that listeners have a perceptual mechanism to deal with uncertainty about and variability in English vowels but not consonants. She claimed that her findings would have implications on approaches to speech perception in both normal and noisy environments, because if the identity of a vowel is inherently more flexible (or perhaps under-defined) than that of consonants, consonants are more likely to be affected by distortion

[5] Earlier studies such as Fletcher (1929), Dudley (1940), and Owens, Talbott, and Schubert (1968) examining related issues such as segment discriminability frequently made several comments on consonants being the major carrier of speech information. However, such studies were not a direct examination of segmental contributions.

than vowels which would affect speech intelligibility. The extent to which the phonemic inventory of English (specifically, its CV-Ratio) affected these results remains an open question. It is possible that languages with a limited set of vowels in their inventory, such as Spanish, may behave differently, and van Ooijen acknowledged this in her paper.

Four years later, in a joint work with Cutler and other colleagues (Cutler et al., 2000), van Ooijen addressed the issue of CV-Ratio in a cross-linguistic comparison study, using the above-described reconstruction task with Dutch and Castilian Spanish. The opposition between Dutch, which has 16 vowels and 19 consonants, and Spanish, which has 5 vowels and 20 consonants, produced a useful comparison between the two languages in terms of CV-Ratio. In their first experiment, a group of native Dutch speakers were presented with disyllabic nonwords. Each nonword had the potential to become a real word if a vowel or consonant was altered; furthermore, in each word, an approximately similar number of vowels and consonants were potential targets for alteration. The three conditions from van Ooijen (1996) – in which participants were free to make any alteration, or were restricted to either VO or CO alterations – were adopted. The findings were consistent with those of van Ooijen (1996); the proportion of correct responses involving vowels was higher than that of consonants, the reaction time was faster in the vowel condition, and the error rate was higher in the consonant condition. In their second experiment, a parallel design was used with Spanish. The outcome was almost identical to the one obtained with Dutch. Interestingly, in the CO condition, a number of vowel intrusions appeared in

both experiments. This was interpreted as evidence that it was easier for the participants to change the vowel, even in the consonant condition where they had been instructed to change only consonants.

Cutler et al. (2000), along with confirming the previous results found in English by van Ooijen (1996), ruled out the possibility that consonant's privilege could be due to CV-Ratio because the results from two languages with very different CV-Ratios, namely Spanish and Dutch, were parallel to one another. One remaining issue is the fact that Dutch, English, and Spanish belong to the same Indo-European language family, and share a similar concatenative morphological system that does not treat vowels and consonants differently. Other languages that demonstrate a special treatment of vowels vs. consonants, such as Arabic and Hebrew, may display a divergent pattern.

Owren and Cardillo (2006) pursued the contribution of vowels vs. consonants to speech intelligibility by adapting the methodologies of Cole et al. (1996) with some modifications. First, silence replacement was used in lieu of noise in order to avoid any differential phonemic restoration effects between consonants and vowels, a concern that was pointed out in the conclusion by Cole et al. (1996). Second, isolated words were used instead of sentences, in order to eliminate the potential contribution of contextual information. The words varied from two syllables to five syllables (from a methodological point of view, note that this allows for significant variation in the ratio of vowels and consonants from one stimulus word to the next).

In a same-different task with three conditions – intact speech, VO, and CO –

the participants judged whether the stimuli they heard were produced by the same or a different speaker and whether they had the same or different meanings. Each pair of words was either clearly similar (i.e., synonymous) or clearly dissimilar. Although beyond of the scope of this study, one portion of their findings illustrated that, for the talker's identity, indexical information is carried by vowels more than by consonants, which was not completely surprising because vowels have long been identified as personal information-bearing elements (see e.g., Cooper et al., 1952; Ladefoged & Broadbent, 1957). The other portion of the findings assert that consonants are more useful for listeners to recognize word meaning. Note that this finding was inconsistent with the results found by Cole et al. (1996), although Owren and Cardillo (2006) used words in isolation while Cole et al. (1996) used words embedded in sentences.

As in earlier studies, there was a concern that coarticulatory cues in vowels had affected the discrimination results; the researchers attempted to eliminate this concern by running an additional experiment. In the VO condition, they replaced 50% of the onset and offset of each vowel with silence. They stated that, if vowels carry some co-articulatory information that helps listeners identify the adjacent consonants, then reducing vowels to this radical extent should eliminate such effects. The results showed a sharp decline in the participants' performance on the meaning discrimination, but not in talker discrimination. This was taken as evidence that consonantal cuing is linguistic while vocalic cuing is indexical.

The results on meaning discrimination conflict with previous findings from

Cole et al. (1996), Burkle (2004), and Burkle et al. (2004). To explain such a discrepancy, Owen and Cardillo (2006) discussed several factors that could have affected the results of Cole et al (1996). First, the sentences used in the study by Cole and colleagues probably contained contextual information that could have allowed listeners to make educated guesses about what word they were hearing. Second, the white noise used to mask segments in these studies may have provided a less effective test of the hypothesis than silence, since listeners are known to "restore" phonemes in noisy conditions (see e.g., Samuel 1981). Third, using noise to replace the segment in its entirety may have resulted in some background masking that affected perception of the neighboring segments. Fourth, notwithstanding the care that Cole et al. took to eliminate any coarticulation effects, a reduction/expansion of 10 ms may have preserved some transitional elements, which were ruled out in Owren and Cardillo (2006) by increasing the magnitude of reduction to 50% of the entire segment. Nevertheless, follow up studies have confirmed that manipulating boundaries to include or eliminate coarticulatory information has an imperceptible or no effect on the contribution of vowels vs. consonants in word and sentence intelligibility, especially for normal hearing listeners (see e.g., Fogerty & Humes, 2012; Fogerty & Kewley-Port, 2009).

Using a different experimental technique, Bonatti, Pena, Nespor and Mehler (2005) conducted a study that investigated the informational role of consonants and vowels using an artificially manipulated language. In a familiarization phase, French participants were informed that they would hear an artificial language, and then

were presented with a stream of speech with different fixed distributions of vowels and consonants. In the later test phase, they were presented with pairs of items and asked to indicate which series of segments sounded like a word in the artificial language. As specified in their experimental design, a word is a set of three consonants such as *b_d_k* concatenated with different patterns of vowels such as *biduka, bidoke, byduka*, and *bydoke*, whereas a nonword (called "part-word" in their study) is a series of consonants and vowels belonging to different words but combined together to form a word-like pattern. This design resembles lexical structure in Semitic languages where vocalic information determines grammatical properties and consonants individuate lexical items. In the second experiment, they reversed the role of vowels and consonants; vowels served as the root-like component. The participants were expected to use the so-called "transitional probabilities" among consonants in one experiment and among vowels in another to identify what is intended to be a word. The findings from their experiments showed that the listeners were successful in using consonants, but not vowels, to extract words over part-words, which was interpreted as evidence that the consonantal tier can play a key role in identifying words in language processing. The preferential role for consonants in speech was not due to a numerical asymmetry in language CV-Ratio leading to favoring consonants over vowels, because the French inventory has a relative balance between consonants (17) and vowels (16). This interesting conclusion was also supported in a subsequent study (Toro, Nespor, Mehler, & Bonatti, 2008).

Thus far, I have discussed the major studies that support the privilege of vowels as well as those that support the privilege of consonants. In the following subsection, I will discuss the studies that show different contributions of vowels and consonants in different contexts.

2.1.3 Vowel vs. Consonant's Privilege

A recent and important study by Fogerty, Kewley-Port, and Humes (2012)[6] sought to demonstrate a distinction between vowel vs. consonant contributions in isolated words vs. sentences, with a focus on older listeners who were either normal-hearing or impaired-hearing. The study also took into consideration the duration of vowels and consonants. It was already established in Fogerty and Humes (2012) and Fogerty and Kewley-Port (2009) that shifting vowel vs. consonant boundaries does not consequently change the segment contribution, however it *does* change the relative length of a given segment, which in turn has some effects on elderly listeners.

Therefore, Fogerty et al. (2012) manipulated the proportion of total duration in some versions of the stimuli that were adopted from their previous study (Fogerty & Kewley-Port, 2009) which involved the selection of some sentences from the TIMIT database and noise replacement of segments. Because the segmental boundaries were already specified in the database by expert phoneticians, the original segmentation was used as a baseline condition where all durations were preserved. However, in

[6] Two other studies preceded this study and reached similar conclusions: Fogerty and Kewley-Port (2009) on segmental contributions in isolated words and Fogerty and Humes (2010) on segmental contributions in sentences. Fogerty et al. (2012), reported here, compares words and sentences and uses different groups of listeners.

other conditions and prior to noise replacement, the proportion of total duration technique manipulated the relative duration of vowels vs. consonants to include more transitional information by shifting consonants towards vowels by 15% or 30%, or shifting the vowels towards consonants by 15%. To illustrate this, when a consonant is 100 ms and a vowel is also 100 ms, shifting the consonant towards the vowel by 15% would make the consonant duration 115 ms. When replacing the segments with silence or noise, the replacement of the consonant would be longer than the original segment, while the replacement of the vowel would be shorter than the original segment. The subjects were instructed to orally repeat the stimuli in the sentence condition, and to type words in the word condition. The chief finding was that vowels carrying high-order cues contribute to intelligibility more than consonants do, but only in sentences and not for words in isolation.

So far, I have reviewed the scholarship on the contribution of vowels vs. consonants in both word recognition and sentence intelligibility based on experiments with data from English, Dutch, Spanish, Mandarin, and an artificial language presented to French speakers. As a rough statement, with differing experimental techniques and tools, these findings suggest that vowels are more important for sentence intelligibility while consonants are more important for word intelligibility. However, one important critique is that all of these previous studies worked on languages that have a similar morphological treatment of vowels and consonants. The difference in contribution of vowels vs. consonants may turn out to be language-specific, an idea that finds some support in the artificial-language results of Bonatti and colleagues (2005)

and Toro et al. (2008). As noted earlier, it is also possible that the relative number and the distribution of vowels vs. consonants in each stimulus item, which was not fully controlled for in previous studies, might have affected the findings. Moreover, it is not clear how the contextual benefits provided in sentence conditions, but not in isolated word conditions, may have affected results. All these potential methodological problems are eliminated in the current study design.

This section has looked at the role of segmental information (vowels vs. consonants) in speech intelligibility, which is the main concern in the present study. Previous studies on word recognition in Semitic languages have shown that listeners decompose words into individual morphemes corresponding to roots and patterns (i.e., sequences of consonants, and sequences of vowels). Given the present study's hypothesis that the morphological role of consonants vs. vowels in Semitic languages is expected to affect (or even override) their phonetic-based role, in the next section I review previous findings on the respective roles of vowels and consonants in word recognition in Semitics.

2.2 Segmental Information and Lexical Recognition in Nonconcatenative Languages

Several psycholinguistic studies have examined whether the root and/or pattern is used as a unit to represent, access, and recognize lexical items in Semitic languages. As we will see, the findings from such studies consistently report that Semitic language speakers heavily rely on the root to represent, access, and recognize items from their mental lexicon.

Many studies that investigated word recognition used "priming", a technique that presents the participant with a visual or auditory stimulus item that is deemed to bear some relationship with a target stimulus (see e.g., Diependaele, Grainger, and Sandra, 2012). For instance, responses to a target word "driver" can be facilitated by presentation of a semantically related prime word, such as "car", or by a morphologically related prime word, such as "drive". The priming and target items can be presented in succession or separated by short or long lags, and the task can be lexical decision, naming, or another type of response as instructed. The priming effect appears as a facilitatory or inhibitory realization in the response, usually the magnitude of reaction time and/or rate of errors.

This technique helps researchers model how words are stored in the mental lexicon. For example, priming experiments have played an important role in the debate about morphological *decomposition* (see e.g., Caramazza, Laudanna, and Romani, 1988; Diependaele et al., 2012; Taft, 1985; Taft & Forster, 1975); that is, do we store morphologically complex words in their full form, or as individual morphemes? For example, do we store the English word "reader" as one holistic unit, "reader", or as two individual morphemes, "read" plus "er"? Even researchers who believe in decomposition subdivide further into two groups: the "full decomposition" group argues that every complex word is decomposed into its constituent morphemes, while the "dual-route" group argues that both whole word and decomposition are simultaneously available. While a large body of evidence supports decomposition, several problems for this account remain unsolved. For example, irregular forms such as "find"

vs. “found” do not always show a clear morphological structure (see e.g., Caramazza, Laudanna, & Romani, 1988). If decomposition occurs by stripping off the affixes and suffixes, then it is unclear how an irregular form is accessed through the base form. Furthermore, non-linear morphology builds lexical items in nonconcatenative manner (see e.g., Deutsch & Frost, 2003; Feldman, 2013; Shimron, 2003). This is problematic for any decomposition-based lexical approach partially because the root is an abstract morpheme that does not form a lexical unit on its own to be accessed, and partially because the root is distributed over and split by the vocalic pattern, which makes stripping one from the other unclear.

Using the priming technique, Bentin and Feldman (1990) sought evidence, from Hebrew, for the contribution of morphological (and semantic) relatedness effects to repetition priming, known as the “morphological repetition effect”. The morphological repetition effect is the facilitation, shown by short reaction time and/or low error rate, that occurs when the prime and target both retain the morphological stem. For example, in English, the word “repay” primes the morphologically related “payment”. This effect was also found in other languages such as Serbo-Croatian (Feldman & Moskovljević, 1987) such that the morphological repetition effect persists in long lags while semantic relatedness decays in short lags. Such findings motivated researchers such as Bentin and Feldman (1990) to examine the relative contribution of morphological vs. semantic relatedness to word recognition in Hebrew as a Semitic language.

Using nonwords and Hebrew words that are morphologically related (i.e., sharing a root), semantically related (i.e., sharing meaning but from different roots),

or morphologically and semantically related (i.e., sharing both a root and similar meaning), the experimenters tested for facilitation in a lexical decision task where the stimuli were presented visually. Table 2 below depicts the different types of relatedness. The root here is “סְפָר” /sfr/, which is shared between the priming words “סִפְרִיָּה” /safria/ “library” and two target items “מספר” /mispar/ “number”, and “סַפְרָן” /safran/ “librarian”. The last word also shares some semantic relatedness with the priming item while the target “קריאה” /kria/ “reading” only shares some semantic relationship with the priming stimulus. Note that the stop consonant /p/ alternates with the fricative /f/ in Hebrew (see e.g., Martinez & Müllner, 2015).

Table 2. Example of root-pattern interdigitating

	Priming	Target		
		Semantic	Morphological	Semantic + Morphological
Hebrew	סִפְרִיָּה	קריאה	מספר	סַפְרָן
Transcription	/sifria/	/kria/	/mispar/	/safran/
Translation	Library	Reading	Number	Librarian

The reaction time as well as the error rate confirmed the results found in other languages; morphological facilitation was maintained over long lags while semantic facilitation was not. However, the root in Semitic languages differs from the stem in other languages such as English, since the Semitic root is just an abstract discontinuous unit, such as /sfr/, while the English stem is usually a word that can stand on

its own, such as “pay”. The authors concluded that the root and pattern are unlikely to have separate lexical units in the lexicon, and suggested a mechanism of “extraction” rather than decomposition.

However, in a later study, Feldman and colleagues (Feldman, Frost, & Pnini, 1995) compared decomposition in English and Hebrew. The necessity for this study came from the realization that, although both English and Hebrew can be categorized as inflected languages, they differ in the manner by which they generate different lexical items from the stem (in English) and root (in Hebrew). In this study, the researchers used the so-called “segment-shifting task” in which the participants were presented with visual materials and asked to segment and shift a certain segment from a source word to a target word and then to name it as quickly as possible. The segmented and shifted element can be a morpheme on its own, such as EN in the English word “harden”, or a mere string of segments, such as EN in the English word “garden”. The participants were instructed to attach this element to another item such as “bright”, which was presented immediately after the -EN bearing item. It was anticipated that facilitatory effects would be greater for the morphemic sequence of letters, such as EN in “harden”, if they are indeed available for the listeners to strip off and use for morphological affixation, as theories of decomposition predict. That is, the morphological affixation of “bright” to become “brighten” was expected to be faster after “harden” than after “garden”. The results from their first two experiments on both real words and pseudowords confirmed the hypothesis with English, and were consistent with Feldman’s previous findings on Serbo-Croatian (Feldman, 1991).

In their experiment on Hebrew, the segment-shifting task took a slightly different form. The participants were presented with written Hebrew source words and asked to detach the pattern and apply it to a pseudoroot (a sequence of three letters that are not used in Hebrew for any meaning). The source items could be either words with a transparent root that could be used to form other words, or words with an opaque root that does not appear in other Hebrew words. The expectation was that in morphological processing, by analogy to English and Serbo-Croatian, detaching and applying the pattern to the pseudoroot would take a shorter time for the transparent root than for the opaque root. The outcome supported their predicted results. The findings support the hypothesis that Hebrew lexical organization is based on morphemic units, i.e., root and pattern, and that the decomposition found in English and Serbo-Croatian is also available in Hebrew irrespective of the difference in the stem or root-affix appending process.

The results form Bentin and Feldman (1990) and Feldman, Frost, and Pnini (1995) combined together show that, in word recognition, root and pattern in Hebrew are equivalent to stem and affix/suffix in English, respectively. This is crucial for the present study because one of our goals is to examine whether the morphological characteristics in Semitic languages, and in Arabic in particular, can affect segmental contributions to speech intelligibility found in other languages. However, two issues are still associated with the two studies in question. Bentin and Feldman (1990) addressed morphological and semantic relatedness, but phonological and/or orthographic relatedness (see Table 3) is another factor that a comprehensive study may

seek to consider, which is what Feldman (2000) explored later. That is, many words that are morphologically related are orthographically related too, and if the target and prime both share the base morpheme, the orthographical effect may confound the results. Note, however, that Feldman, Frost, and Pnini (1995) and many other studies used visual presentation of stimuli, which may have introduced an orthographic interference that would not occur with auditory presentation.

Table 3. Example of semantic vs. morphological vs. orthographical relatedness

Relation	Priming	Target		
		Semantic	Morphological	Orthographic
English items	Vow	Pledge	Vowed	Vowel

Feldman (2000) investigated the morphological effects relative to semantic and orthographic effects in English and not with any Semitic language, but her study is important as it complements the study by Bentin and Feldman (1990) and provides a preview of another study on Hebrew, which will be introduced shortly. Feldman (2000) used prime and target words that, as shown in Table 3, were morphologically related (vowed-vow), orthographically related (vowel-vow), or semantically related (pledge-vow) in an immediate vs. long-term lexical decision task. Based on the patterns they found for latencies and accuracy rates, this author concluded that morphological priming effects are distinct from those contributed just by meaning alone or just by orthography alone, especially in long lags. The series of experiments re-em-

phasized the role of morphemic units in word recognition and controlled for potentially confounding effects from orthography/phonology.

Deutsch, Frost, Pollatsek, and Rayner (2000) contrasted morphological relatedness with orthographic relatedness to address the factors that affect word identification in Hebrew. They purposely used a new technique known as "parafoveal preview benefit" to test whether it would lead to the same or different findings that previous techniques have led to. The parafoveal preview benefit technique can be explained as following:

> The parafoveal target word is initially changed by replacing it with, for example, a semantically or orthographically similar word. An invisible boundary is placed to the left of the parafoveal word, and when the eyes cross the boundary, the parafoveal word is changed to its intended form. The display change is made during the saccadic eye movement, and since vision is greatly suppressed during saccades, the reader does not perceive the actual change taking place (Hyönä, Bertram, & Pollatsek, 2004, p. 524).

The authors used three types of stimuli: identical, morphologically related, and orthographically related Hebrew prime and target words (refer to Table 4). The participants were instructed to name the target as rapidly as they could, and the reaction time and error rate were calculated to examine preview benefit effects. It was found that morphological facilitation was greater than the facilitation in the orthographic condition, which also supports the findings from Feldman (2000) but with a new technique.

Table 4. Example of identical vs. morphological vs. orthographical relatedness

	Priming	Target		
Relation		Identical	Morphological	Orthographic
Items	גִּדּוּף	גִּדּוּף	**גדף	**גיף
Transcription	/giduf/	/giduf/	/gdp/	/gyp/
Translation	Cursing	Cursing	~	~

Note 1: The asterisk (*) indicates missing letters. The translation is not available because the items are not complete words.

There remains a central distinction in relation to the difference between root and pattern as two morphemes and between verbal and nominal items in Hebrew. Although it makes sense that most studies looked at the root as a morphemic unit that is comparable to the stem in concatenated languages, the pattern should still be contemplated. Deutsch and Frost (2003) reviewed many studies (including theirs) on lexical access and organization in nonconcatenative morphology and reported that, while root can facilitate both verbs and nouns, pattern only facilitates verbs. They developed a dual-route model for the Hebrew lexicon that allows both whole word and sub-word, which may apply to Arabic as a Semitic language, although nominal and adjective forms in Arabic are more constrained by the root and pattern than in Hebrew (see e.g., Abu-Rabia, 1997; Holes, 2004; Mimouni, Kehayia, & Jerema, 1998; and Ravid & Shimron, 2003). In sub-word retrieval, verbal items are retrieved either by root or pattern while nominal items are only retrieved by root. Figure 2 below illustrates this architecture.

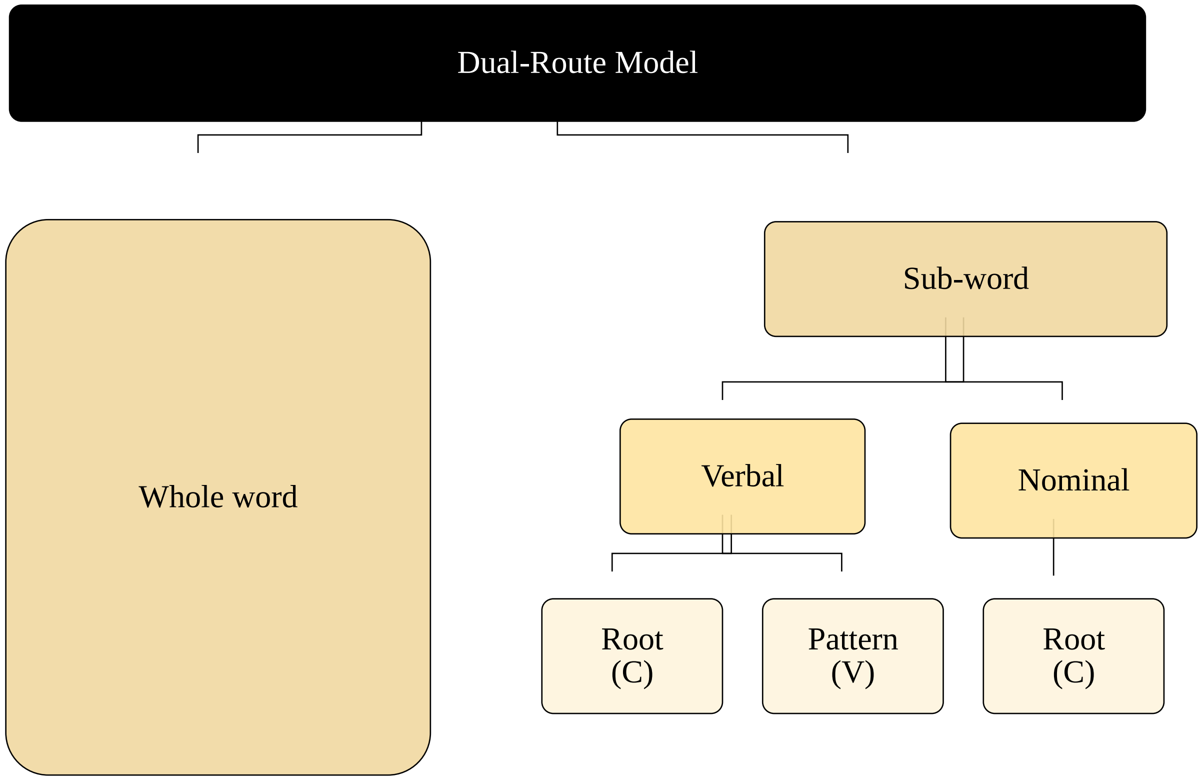


Figure 2. Representation of the dual-route model for root-pattern lexicon

Thus far, I have reviewed the most relevant studies on root-pattern morphological processing and word recognition.[7] At this point, the necessity for an experimental work to contrast the phonetic information inherent in vowels and consonants with the morphological information they carry according to the requirements of the root-based system as well as the relative distribution of segments has become more obvious. The next chapter will present the first experiment on segmental contributions in Arabic.

[7] Other studies about language processing in Semitic/root-based morphology languages can be found in Shimron (2003).

⇐ Chapter Two

⇒ Chapter Four

CHAPTER THREE:

SEGMENTAL CONTRIBUTIONS AND RATIOS IN ARABIC

As reviewed in previous chapters, researchers have investigated the contribution of segment type (consonants and vowels) to speech intelligibility (words and sentences) in both normal-hearing and hearing-impaired speakers of languages such as English (Cole et al., 1996; Kewley-Port et al., 2007), Spanish, and Dutch (Cutler et al., 2000). The main findings show that, while consonants play a bigger role in the intelligibility of isolated words in these languages, vowels contribute more to the intelligibility of complete sentences. This conclusion emerged from experiments on concatenative languages, whose morphological systems do not functionally distinguish between vowels and consonants. However, the contributions of segment type to speech intelligibility may not be solely based on phonetic properties, but also on other factors such as morphological characteristics. Specifically, since Arabic employs consonants as primary carriers of lexical information, we can expect consonantal information to play a greater role in speech intelligibility as listeners must rely on such information to recognize, access, and retrieve lexical items and consequently comprehend speech. It thus follows that the magnitude of consonantal information present in, or absent from, speech would substantially affect speech intelligibility.

In this Chapter, I seek to answer two major questions about the roles of consonantal and vocalic information and the ratio of their occurrence in speech with respect to speech intelligibility:

- *What are the segmental (vocalic vs. consonantal) contributions to speech (words vs. sentences) intelligibility in nonconcatenative languages, specifically in Arabic?*
- *How does segmental ratio (balanced vs. imbalanced) affect speech (words vs. sentences) intelligibility in nonconcatenative languages, specifically in Arabic?*

To answer these questions, I conducted three experiments in which segmental contributions and ratios were examined at both word and sentence levels in Arabic. Experiment 1 was designed primarily to investigate individual words. Experiment 2 was designed to test individual words with different S-Ratios embedded in sentences that had a balanced S-Ratio, and to test ratio-balanced sentences. Experiment 3 investigated full sentences with an imbalanced S-Ratio.

3.1 Experiment 1

In this experiment, I investigated segmental contributions and ratio at the word level in Arabic, using the silence replacement method. Given the importance of consonants in the Arabic morphological system, and given the results of previous studies on speech intelligibility that have emphasized the role of consonants even in non-Semitic languages, I began with a straightforward prediction that consonants should play a greater role in the intelligibility of Arabic words than vowels do. In addition, I made the prediction that S-Ratio would reveal a novel finding. Specifically, a word with a 3C-2V ratio, for example, would be more intelligible than a word with a 2C-3V ratio. The effect of S-Ratio is an aspect previous studies have overlooked and may be particular to Arabic and other nonconcatenative languages.

3.1.1 Methodology

The methodology used in this experiment is laid out in detail below. The subsequent experiments from Chapters 3 and 4 will follow approximately the same methodological approach; any differences will be stated and justified.

3.1.1.1 Overview. I composed a list of 24 MSA words, both nominal and verbal, of similar frequencies (see Table 5 and Appendices A and B). Each word consisted of either a balanced or imbalanced number of vowels and consonants, with ratios 2C-3V, 3C-3V, or 3C-2V as in /daʕaa/[8] "called", /taabaʕ/ "followed up", and /samiʕ/ "heard", respectively. After the stimuli were recorded by a native Arabic speaker, I replaced consonants (for the VO condition) or vowels (for the CO condition) with silence lasting the exact duration of the original segment. For instance, the word /taabaʕ/ would be /t~~aa~~b~~a~~ʕ/ in the CO condition and /~~t~~aa~~b~~a~~ʕ~~/ in the VO condition, where the strikethrough represents silence replacement. An additional difference in the current study design from that of other studies is the distinction between verbs and nouns. It has been claimed in the literature that word access in Semitic languages takes place by the root (i.e., consonants) and pattern (i.e., vowels) in verbs and by the root in nouns (Deutsch & Frost, 2003), although the outcome of the current study did not show a difference in speech intelligibility based on this distinction. This design resulted in three types of comparisons: CO vs. VO, 2C-3V vs. 3C-3V vs. 3C-2V, nouns

[8] Note that here I adopt the view that long vowels take two slots and that long vowels in most cases are a result of phonological processes in which the feature [+ Long] added one identical short vowel to make the two adjacent vowels appear as one long segment. For a discussion on this, refer to Brame (1970) and Sevald (1996).

vs. verbs (N-V). I presented twenty Arabic speakers, of different dialects, with auditory stimuli and asked them to record what they could hear. I then scored their responses, assigning 1 point to responses that correctly recognized the word and zero points to responses that did not record the word as was exactly presented.

3.1.1.2 Word length. Each word consisted of an average of 5.33 segments. Because some words had long vowels that had almost double the duration of a short vowel, I counted each long vowel as consisting of two segments. Geminate consonants were also counted as two segments, regardless of whether they were true or fake. Note that both long vowels and geminates are phonemic in Arabic, contrasting with short vowels and singleton consonants, respectively. A third of the stimuli consisted of 3 consonants and 2 vowels (5 segments [3C-2V]), one third consisted of 3 vowels and 3 consonants (6 segments [3C-3V]), and one third consisted of 2 vowels and 3 consonants (5 segments [2C-3V]).

3.1.1.3 Word selection and frequency. I chose MSA words that have very frequent patterns and frequently appear in daily use, not only in MSA but also in dialectal Arabic (with minor changes in phonetic realizations). I then checked the word's frequency on the Arabic Corpus Tool, arabicorpus.byu.edu (Parkinson, 2006), to make sure that the words are not only frequent but also that they have similar frequencies across all conditions (refer to Table 5 below).

Table 5. Word frequencies in Experiment 1

	S-Ratio		
	2C-3V	3C-3V	3C-2V
Mean	36.06[9]	35.65	33.23
STD	8.36	10.60	11.44

3.1.1.4 Judgements. All words were submitted to 10 native speakers of Arabic to rate, on a 1 to 7 scale, their familiarity with each word. The list was sent to the five judges in written form once, and later to five judges to rate the words in spoken form. Any word that was overall rated less than 6 was replaced, and a final version of the stimuli passed the criteria with an average of 6.95 in written form and 6.96 in spoken form (see Appendix B for the sentence ratings).

3.1.1.5 Recording and stimulus preparation. The words were recorded by a male native speaker of Arabic with a linguistics background. Although previous studies used stimuli produced by a variety of different speakers, I determined that having only one speaker would rule out any possibility of inter-speaker variation effects. The speaker was asked to record the sentences on Praat (Boersma, Paul, Weenink, David, 2017) in a careful manner to make the recording as clean, clear, well-articulated, and noise-free as possible. This recording helped make the boundaries of segments easily identifiable in the process of silence replacement. It also helped in

[9] It should be noted that the entire corpus was not used for this frequency determination; rather, the corpus was set to contain only newspapers in order to exclude some texts that are not suitable for the current study, such as medieval literature, Islamic traditions, dialectal Arabic, and texts for Arabic learners.

allowing some segments that a speaker may not fully release at the end of the sentence to be produced as a full segment, such as /h/ in /xatˤiir-ah/ "dangerous-F". The speaker was also asked to record the words at an intermediate speech rate and within a single session to ensure that each word lasted for a similar length of time. This also prevented any reduction due to fast speech as it was important to keep the number of vowels and consonants equal and to ensure that each one of them was fully articulated.

I then annotated each word and marked the boundaries of vowels and consonants segment by segment. In marking the boundaries of stops, both closure and burst were included in the stop duration. For voiced stops, the voice bar appearing in the spectrogram was used as evidence for the stop-vowel distinction and was included in the stop duration. For voiceless stops, the absence of F1 was taken as evidence for the short aspiration noise that follows the release burst and precedes the following vowel, and was included in the consonant duration. The onset of voicing was marked as the beginning of the following vowels.

One problem arose with words that had an initial stop, such as /taabaʕ/ "followed up". Waveforms and spectrograms do not provide definitive information about the onset of such stops when they occur in absolute phrase-initial position, as they did here, and there were two possible solutions. One was to look at the same word in final position in the sentence stimuli prepared for Experiment 2, where it was preceded by a vowel and to use the same closure duration for the same word in isolation. The other solution was to calculate the duration of that particular segment across all

words in the stimulus set, and use the average. The latter solution was preferred since a final position in sentences may not mirror words in isolation, as final position is subject to lengthening effects.

The boundaries of two of the nasal consonants, /n/ and /m/, were sometimes hard to distinguish, especially when occurring pre- or post-vocalically, but I applied the following criteria in all cases. Nasals were identified by the relative lack of aperiodicity, simple waveform pattern, and low frequency spectral components. When a nasal was preceded by a vowel, the sharply-down-pointing formants on the spectrogram were used as an indication of the nasal onset, and the sudden increase in intensity at the beginning of the following vowel was taken as a marker for the ending of the vowel.

Fricatives were relatively easy to identify and mark. The beginning and ending of the fricative noise or turbulence were identified as the onset and offset. Vowels that occurred after fricatives were marked at the point where the higher formants started to become apparent on the spectrogram, while vowels that occurred before fricatives were marked by a drop in intensity and loss of energy, especially in higher frequencies.

For the approximants, the decision on where to mark the boundaries was based on the abrupt change of intensity in F2 and F3 noticeable in the spectrogram, coupled with a dramatic change of the overall amplitude in the waveform. To identify the liquids and glides boundaries as well as the preceding or following vowels, I relied on the liquid/glide-vowel transition of formants where F2 and F3 start to show some

signs of movement and the stabilization points where the formants start to maintain their shape. I made the decision to include the transitions into the vowels rather than the consonants; this was dictated by the hypothesis because, as Wright and Nichols (2014) point out, the experimenter should make decisions that introduce as little bias as possible to the research. The current study argues for the superiority of consonants, and therefore, it was necessary to avoid any possible factor that, albeit minimally, could potentially affect the results. One important consideration was how to deal with the Arabic /r/. In previous studies, the English /ɹ/ was problematic and researchers had to decide whether to consider it as a separate consonant or as part of vowels (*r*-colored). The Arabic /r/, however, is a trill that is, in most cases, clearly visible on the waveform and spectrogram as a triple of the same segment (as shown in Appendix C), and therefore, I considered it as a separate consonant.

The uncrowded space of Arabic vowels, the presence of long vowels, and the careful speech that the talker produced made it a relatively easy task to segment vowels. Difficulties in marking the boundaries arose when the adjacent consonants were approximants, but the decisions on this were reported above.

After marking the segment boundaries for all words, I produced two versions of each word, one in which vowels were replaced by silence and the other in which consonants were replaced by silence, a technique which was has been used by several other researchers (e.g., Owren & Cardillo, 2006). Figure 3 (adopted with a slight modification from Owren & Cardillo, 2006, p. 1730) illustrates segment replacement in the CO vs. VO conditions. To replace each marked segment with a silence, I used the

Edit function on Praat, and chose *Set Selection To Zero*, which turned each selected segment into a silence that had exactly the same duration as the original segment. The vowels and consonants had similar mean durations (see Appendix D for all the duration means).

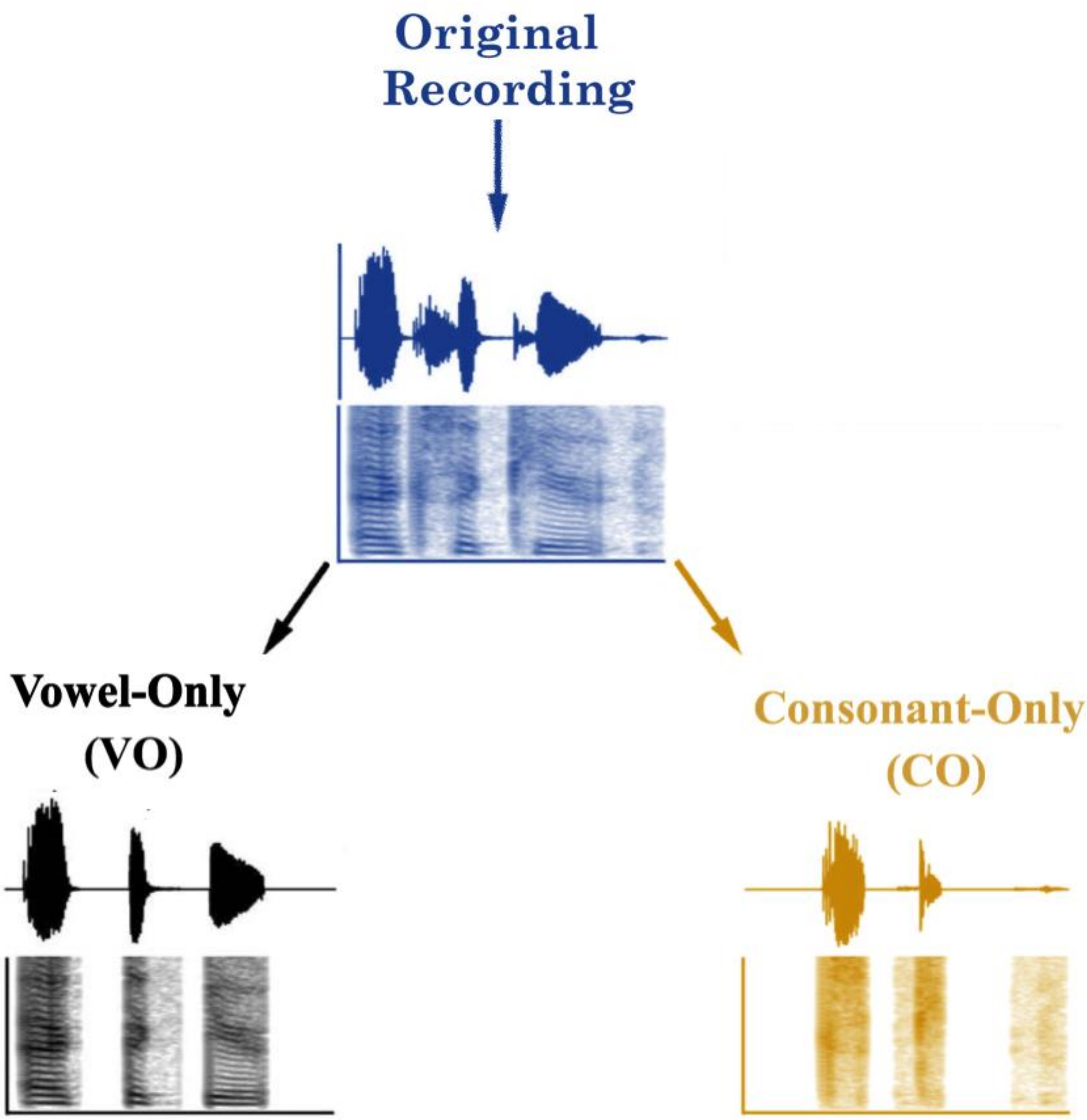


Figure 3. Illustration of segment replacement in CO vs. VO

One potential problem involved substituting initial segments with silence, as it was very likely that the participants would miss the intended beginning of each trial in such cases, thereby affecting the results. To solve this issue, I added a pure tone *stimulus carrier*. The tone length was made to equal double the mean of all segments (see Appendix D for C, V, and C+V durations in the stimuli) in the entire set of stimuli ([M = 152 ms] * 2 = 304 ms), which gave the subject sufficient time to expect the new stimulus and then realize when it had started. The stimulus carrier was also added at the end of each stimulus since some of the stimuli included a silence replacement in the last segment. After all segment-replaced words were saved in a separate file, the amplitude of the audio files was normalized. It was important *not* to normalize the tone amplitude; otherwise, it would have been too loud compared to human speech. This is exactly what Kewley-Port et al. (2007) did with regard to noise replacement; they normalized the 43 sentences but not the noise replacement.

3.1.1.6 Task. In similar previous studies, the subjects were presented with the stimuli and asked to type or write down as much content as accurately as possible. In the current experiments, the task was slightly different due to the nature of the Arabic orthographical system, which does not explicitly represent most vowels. The participants were provided with a high-quality recorder and a microphone, and asked to record what they could hear immediately after they were presented with each trial twice (see Appendix E for user interfaces).

3.1.1.7 Participants. Twenty native Arabic speakers volunteered to participate in the study. The participants come from different Arab countries (Saudi Arabia,

Jordan, Iraq, Syria, and Morocco) and speak different dialects, such as Najdi, Hijazi, Southern Saudi, Jordanian, Syrian, Iraqi, and Moroccan Arabic. All are students (mostly graduate) at the University of Wisconsin-Milwaukee, Cardinal Stretch University, Marquette University, and Concordia University-Wisconsin. Most of the subjects were males (15 males and 5 females), and their ages ranged between 26 and 40 (*M* = 31). They do not speak languages other than Arabic as a native language; English is their second language, except for one subject who speaks French in addition to Arabic and English. Some participants are Linguistics students at the University of Wisconsin-Milwaukee, but none of them had exposure to the stimuli during any phase of the word stimuli preparation and rating.

3.1.1.8 Procedure. The participants were met in the Phonology Lab at the University of Wisconsin-Milwaukee where the experiment took place. Each participant was introduced to a set of instructions, and then before the actual experiment started, the experimenter left the room in order to give the subjects more freedom to utter the words as they wished. No two words from the same condition were adjacent in the stimuli presentation nor did any participant listen to the same word in two different conditions.

Following Kewley-Port et al. (2007), I presented each participant with 6 familiarization trials, that were later excluded from the results analysis. I also followed Kewley-Port et al. (2007) in allowing only two presentations of each stimulus before responding, although some researchers such as Cole et al. (1996) allowed 5 presentations of the same stimulus. In addition, they were instructed to provide their final

response after the second stimulus presentation if they accidentally provided their responses after the first presentation. I also made sure that they followed the instructions by matching the clicking times (the *repeat* and *move on* buttons) to the recording times, and I only found four instances where participants recorded after they listened once. In two of them, the participants reported their responses again after they listened to the stimuli for the second time; in one of them, the participant repeated the same response after they listened for the second time; and in one of them, the participant did not report anything after the second time. Finally, the participants were allowed to set up the volume at a comfortable level and navigate through the experiment at their pace.

3.1.1.9 Data scoring and analysis. Words were assigned one point if the entire word was correctly identified and zero points otherwise. Following previous researchers such as Kewley-Port et al. (2007), small errors such as vowel length contrast were not tolerated and were assigned zero points even if the semantics of the words remained similar.

The dependent variable was binary (0, 1) and therefore, as described below, we conducted statistical analysis using a logistic regression model, which is most appropriate for binary data. In addition, for the purpose of illustrating the general patterns in the results as well as comparing them with previous studies, I also calculated two additional values, a) the overall proportions of intelligibility, e.g., 68% percent for Subject 1, and b) rationalized arcsine units (RAU), which are described briefly in the paragraphs that follow. Note that the proportions and the RAU values were used only

for descriptive purposes. My statistical analysis relied instead on binary (0, 1) values, which has the advantage of retaining every single data point in the results, rather than averaging the results per participant or item.

In speech science studies, researchers calculate the proportion of intelligibility per subject or item and then transform the outputs into angular values. The angular (or arcsine[10]) transformation converts proportional scores into radians to make them more suitable for parametric statistical analysis (Sherbecoe & Studebaker, 2004). However, the output can be difficult to interpret; therefore, two solutions were proposed in the literature to present results in units that can be intuitively interpretable. First, the angular scores can be transformed back to percentages as some researchers in speech science have done (see e.g., Studebaker, Bisset, Van Ort, & Hoffnung, 1982). This solution has some potential problems such as the slight diversion between the original percentage and the transformed percentage. Second, some researchers (Studebaker, McDaniel, & Sherbecoe, 1995) proposed to transform the arcsine values to percentage-like units known as *the rationalized arcsine units* (*RAU*). We used the RAU transformation equations from Sherbecoe and Studebaker (2004) and Studebaker et al. (1985; 1995). Equation 1 below displays the equations executed to obtain the rationalized arsine values.

[10] I am using "angular" instead of arcsine to distinguish it from the rationalized arcsine.

Equation 1. Angular (Arcsine) vs. RAU transformation

$$(1)\quad \text{Arcsine} = 2 * \text{asin}\sqrt{\frac{CR}{N \text{ when } N > 150}}$$

$$(2)\quad \text{Arcsine} = \text{asin}\sqrt{\frac{s}{N \text{ when } N < 150 + 1}} + \text{asin}\sqrt{\frac{s+1}{N \text{ when } N < 150 + 1}}$$

$$(3)\quad \text{Rationalized Arcsine} = \left(\frac{146}{\pi}\right) * \text{asin} - 23$$

CR = Correct responses N = Number of trials

3.1.2 Results

The descriptive results indicate that words were more intelligible in the CO condition (65.00% and 64.03 in RAU) than in the VO condition (7.50%, and 3.38 in RAU). Table 6 shows the overall results (see Appendix F for illustrations).

Table 6. Intelligibility (% and RAU) of CO vs. VO in Experiment 1

CO vs. VO	Responses	Overall (%)	Overall (RAU)
CO	240	65.00%	64.03
VO	240	7.50%	3.38

Figure 4 shows the intelligibility proportions in the CO vs. VO conditions when broken down by S-Ratios. If we look at intelligibility rates across the CO condition, we see an upward pattern such that the more consonants and fewer vowels the stimuli have, the better participants' performance becomes: 82.50% (81.94 RAU) in 3C-2V, 65% (63.81 RAU) in 3C-3V, and 47.50% (47.83 RAU) in 2C-3V. There is no clear pattern in the VO condition, perhaps because the proportions themselves are so

small. Intelligibility proportions range from 5.00% (0.08 RAU) in 3C-2V, dropping to 2.50% (-5.36 RAU) in 3C-3V, while increasing to 15% (15.04 RAU) in 2C-3V. The N vs. V contrast did not show any observable difference and hence will not be discussed here (see Appendix F for rates and visual representations).

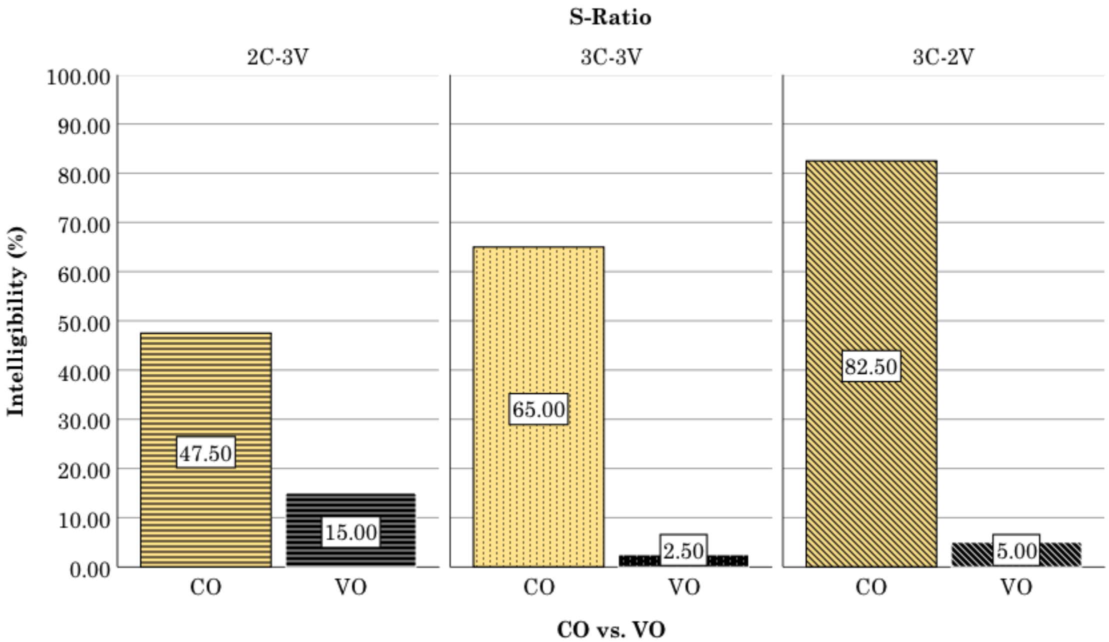


Figure 4. Intelligibility (%) broken down by S-Ratio and C0 vs. VO in Experiment 1[11]

In order to determine if the patterns visualized above are statistically supported, I used a repeated-measures logistic regression model to test for the main effects of and interactions between predictors. As outlined in the methodology section, the design contrasts CO vs. VO, N vs. V, and S-Ratios as predictor[12] variables. The

[11] Note that the error bar is not available on this graph and some other ones in which the dependent variable is binary, but the overall intelligibility was manually calculated.
[12] *Predictor*, *explanatory*, and *independent* will be used interchangeably to describe factors such as VO vs. CO throughout the dissertation regardless of the specific type of the statistical test being used.

intelligibility score is a dichotomous dependent variable with two categories (0 = unintelligible or 1= intelligible). A binominal logistic regression model could have been performed to predict the impact of the explanatory variables on the dependent variable category. However, given that multiple observations were obtained from each individual participant and thus the observations were dependent or correlated, such a model would violate the *independence assumption*. As an alternative, using Generalized Estimating Equations Model (GEE), I fit a repeated-measures logistic regression model with three predictor variables and one binomial dependent variable to test for the main effects of and interactions between predictors (Table 7 summarizes the model information[13]). Subject and trial were specified as a subject variable and repeated effect variable, respectively.

Table 7. Summary of model information in Experiment 1

Model Information	Selection
Dependent Variable	Intelligibility [a]
Probability Distribution	Binomial
Link Function	Logit
Subject Effect and Repeated Effect	Subject and Trial
Working Correlation Matrix Structure	Exchangeable

a. *The procedure models Unintelligible as the response, treating Intelligible as the reference category.*

[13] The subsequent GEE models will be very similar to the current model. Hence, the model information for the next experiments can be found only in the appendices.

A working correlation matrix was set up as *exchangeable* because the goodness of fit showed that it was the best (Quasi Likelihood under Independence Model Criterion [QIC] = 223.65 and Corrected Quasi Likelihood under Independence Model Criterion [QICC] = 221.37). Other options such as *unstructured*, *AR*, and *M-dependent* all produced larger QIC and QICC values. Nevertheless, it should be pointed out that although an *exchangeable* structure was used, all other options produced almost identical parameter estimates (see Appendix G for the GEE information). Because GEE uses iterations, even if the correlation structure was not set correctly the model will ultimately re-set the structure to the right option. The model will still provide a goodness of fit, but the results will ultimately be identical.

The model was able to predict an overall 78.80% of the responses (189 cases) correctly, while 21.2% of the responses (51 cases) were misclassified (see Appendix G). The model showed that CO vs. VO statistically significantly predicted whether speech would be in either the intelligible or unintelligible category (Table 8), Wald $\chi^2(1) = 39.98$, $p = 0.001$. N vs. V did not statistically significantly predict whether speech will be intelligible or unintelligible, Wald $\chi^2(1) = 3.78$, $p < 0.052$, nor did S-Ratio, Wald $\chi^2(2) = 1.70$, $p = 0.42$. The interaction between S-Ratio and CO vs. VO was statistically significant, indicating that the intelligibility of speech with different S-Ratios depended on whether it was presented in the CO or VO condition, Wald $\chi^2(2) = 10.37$, $p = 0.01$. None of the other interactions were statistically significant; CO vs. VO x N vs. V, Wald $\chi^2(1) = 2.225$, $p < 0.136$, and N vs. V x S-Ratio, Wald $\chi^2(2) = 5.276$, $p = 0.072$.

Table 8. Summary of tests of the model effects in Experiment 1

Source	Wald χ^2	*df*	Sig.
(Intercept)	12.101	1	**.001**
CO vs. VO	39.984	1	**.000**
N vs. V	3.782	1	.052
S-Ratio	1.708	2	.426
CO vs. VO x N vs. V	2.225	1	.136
CO vs. VO x S-Ratio	10.372	2	**.006**
N vs. V x S-Ratio	5.276	2	.071

Dependent variable: Intelligibility

Model: (Intercept), CO vs. VO, N vs. V, S-Ratio, CO vs. VO x N vs. V, CO vs. VO x S-Ratio, N vs. V x S-Ratio

I will now take a closer look at the parameter estimates and interpret the odds ratios[14], which are the exponential parameters estimates (Exp(B)) provided in Tables 9 and 10. It is important to note that *intelligible* was set up as the reference category. As such, the predicted probability was membership for *intelligible* rather than *unintelligible*. To better understand the likelihood when Exp(B)<1 and/or b< 0, Table 10 depicts the odds ratios using *unintelligible* as the reference category, and computed proportions out of Exp(B).

[14] Odds ratio is used interchangeably with (or converted to) the so-called relative risk (RR) in speech science and hearing, communication disorders, and speech-language therapy literature to make it easier and more intuitive for readers. For more about odds ratio vs. relative risk, see Osborne (2006).

As the model tested every possible contrast (e.g., CO x N and VO x N), the individual parameter estimates show 7 main effects and 19 two-way interaction effects (see Appendix H for complete tables), but the model sets to zero any interaction with a non-reference category. As shown in Tables 9 and 10, the first notable finding in the main effect portion is that speech in the CO category was significantly predicted to be in the intelligible category with a 131.30 times (≈ 99.24%) higher occurrence than if it was in the VO category, b= 4.878, SE= 0.8421, 95% CI = [3.22, 6.52], Wald $\chi^2(1)$ = 33.54, p= 0.001, Exp(B)= 131.30 (≈ 99.24%), 95% CI = [25.20, 684.07]. In other words, as Table 10 shows, speech in the CO category was significantly predicted to not be in the unintelligible category, and the likelihood of being in that category was only 0.008 times (≈ 00.007%) more than if it was in the VO category, b= -4.878, SE= 0.8421, 95% CI = [-6.528, -3.227], Wald $\chi^2(1)$ = 33.54, p= 0.001, Exp(B)= 0.008 (≈ 00.007%), 95% CI = [0.001, 0.40]. Note that the range between the lower and upper bounds of CI = [0.001 vs. 0.40] is very small, indicating a high odds ratio precision (see e.g., Szumilas, 2010). Speech from the N category was 3.765 (≈ 79.00%) more likely to be in the intelligible category than that from the V category, b= 1.326, SE= 0.5859, 95% CI = [0.177, 2.474], Wald $\chi^2(1)$ = 5.119, p= 0.024, Exp(B)= 3.765 (≈ 79.00%), 95% CI = [1.194, 11.872]. Although the model did not show a statistically significant effect for the S-Ratio explanatory factor in general, speech with 3C-2V S-Ratio in particular was significantly predicted to fall in the intelligible category 11.200 times (≈ 91.80%) more than speech from other S-Ratio categories, b= 2.416,

SE= 0.9431,95% CI = [0.568, 4.264], Wald $\chi^2(1) = 6.563$, *p*= 0.01, Exp(B)= 11.200 (≈ 91.80%), 95% CI = [1.764, 71.121].

When it comes to factor interactions, we only have two statistically significant interactions to report, both involving the 3C-2V S-Ratio. Because both outputs have Exp(B)<1, we can use Table 10 to interpret odds ratios. The model significantly predicted that the likelihood of speech with a 3C-2V S-Ratio being in the unintelligible (rather than intelligible) category was dependent on being in the CO category 16.654 times (≈ 94.33%) more than in other categories, *b*= 2.813, *SE* = 0.9887, 95% CI = [0.875, 4.751], Wald $\chi^2(1) = 8.092$, *p*= 0.004, Exp(B)= 16.654 (≈ 94.33%), 95% CI = [2.398, 115.648]. The model also significantly predicted that the likelihood of speech with a 3C-2V S-Ratio being in the unintelligible (rather than intelligible) category depended on being in the V category 1.949 times (≈ 66.09%) more than in other categories, *b*= 0.667, *SE* = 0.3204, 95% CI = [0.039, 1.295], Wald $\chi^2(1) = 4.337$, *p*= 0.037, Exp(B)= 1.949 (≈ 66.09%), 95% CI = [1.040, 3.651]. It should be noted that the interaction between CO and 3C-2V is more robust (odds ratio = 8.092, and *p*= 0.004) than that between V and 3C-2V (odds ratio =1.946, and *p*= 0.037), which is consistent with the observation that the CO vs. VO x S-Ratio interaction was statistically significant in the main model ($p<00.006$) while the N vs. V x S-Ratio interaction was not ($p=0.072$).

In summary, the model shows that segment type and segment ratio were two major explanatory factors that contributed to the data fit both for their main effects and interaction effects. The odds ratios for main effects and interaction effects also showed high values (no less than 90% after their transformation to proportions) for both factors.

Table 9. Summary of the GEE parameters from Experiment 1 (Ref. Category= Intelligible)

Parameter [a]	B	Std. Error	95% Wald (CI)		Hypothesis Test			Exp(B)	95% Wald CI for Exp(B)	
			Lower	Upper	Wald χ^2	*df*	Sig.		Lower	Upper
(Intercept)	-4.513	.7595	-6.002	-3.024	35.311	1	**.000**	.011	.002	.049
[CO-VO=1[b]]	4.878	.8421	3.227	6.528	33.546	1	**.000**	131.305	25.203	684.076
[N-V=1]	1.326	.5859	.177	2.474	5.119	1	**.024**	3.765	1.194	11.872
[S-Ratio=3C-2V]	2.416	.9431	.568	4.264	6.563	1	**.010**	11.200	1.764	71.121
[CO-VO=1] x [S-Ratio=3C-2V]	-2.813	.9887	-4.751	-.875	8.092	1	**.004**	.060	.009	.417
[N-V=0] x [S-Ratio=3C-2V]	-.667	.3204	-1.295	-.039	4.337	1	**.037**	.513	.274	.962

Dependent variable: Word Intelligibility

Model: (Intercept), CO vs. VO, N vs. V, S-Ratio, CO vs. VO x N vs. V, CO vs. VO x S-Ratio, N vs. V x S-Ratio

a. Reference category is: *Intelligible.*

b. CO is coded as 1, VO as 0, N as 1, and V as 0.

Table 10. Summary of the GEE parameters from Experiment 1 (Ref. Category= Unintelligible)

Parameter [a]	B	Std. Error	95% Wald CI		Hypothesis Test			Exp(B)	95% Wald (CI) for Exp(B)	
			Lower	Upper	Wald χ^2	*df*	Sig.		Lower	Upper
(Intercept)	4.513	.7595	-3.024	6.002	35.311	1	**.000**	91.193	20.583	404.035
[CO-VO=1[b]]	-4.878	.8421	-6.528	-3.227	33.546	1	**.000**	.008	.001	.040
[N-V=1]	-1.326	.5859	-2.474	-.177	5.119	1	**.024**	.266	.084	.837
[S-Ratio=3C-2V]	-2.416	.9431	-4.264	-.568	6.563	1	**.010**	.089	.014	.567
[CO-VO=1] x [S-Ratio=3C-2V]	2.813	.9887	.875	4.751	8.092	1	**.004**	16.654	2.398	115.648
[N-V=0] x [S-Ratio=3C-2V]	.667	.3204	.039	1.295	4.337	1	**.037**	1.949	1.040	3.651

Dependent variable: Word Intelligibility

Model: (Intercept), CO vs. VO, N vs. V, S-Ratio, CO vs. VO x N vs. V, CO vs. VO x S-Ratio, N vs. V x S-Ratio

a. Reference category is *Unintelligible.*

b. CO is coded as 1, VO as 0, N as 1, and V as 0.

3.1.3 Discussion and Conclusion

Experiment 1 was designed to examine the contributions and effects of segment type (CO vs. VO) and S-Ratios on speech intelligibility in isolated Arabic words. The results show that consonants exhibited a greater contribution to speech intelligibility. The association of speech with either the CO category or the VO category was a statistically significant predictor of whether speech was intelligible or unintelligible. Specifically, speech that preserved consonantal information was more likely to be in the intelligible category than speech that preserved only vocalic information. This consonantal superiority with respect to speech intelligibility at the word level in Arabic supports previous findings in the literature. More importantly, the interaction between S-Ratio and CO vs. VO was also a significant predictor of speech intelligibility, demonstrating that participants were better able to recognize isolated words that contained a higher ratio of consonants to vowels. This emphasizes the fact that S-Ratio is an important factor in speech intelligibility and, furthermore, suggests that segmental contributions to speech intelligibility are tied to S-Ratios. In what follows, I discuss each of these findings in more detail.

3.1.3.1 Segmental contributions. In isolated words, participants were better able to recognize the intended words when consonants alone were present compared to when vowels alone were present. This pattern is almost the reverse of what Cole et al. (1996) found in their study on English, in which performance was better in the VO condition. One key difference between the current experiment and that in Cole et al. (1996) is that Cole et al. assessed word intelligibility within the context of complete

sentences, rather than with isolated words. That is, they presented participants with sentences and calculated accuracy rates for both individual words within the sentence as well as for the sentence as a whole. Since words were not presented in isolation, it could be the case that the participants benefited from the sentence-level prosodic and suprasegmental cues, absent in the isolated word condition of the current experiment. This was alluded to by Owren and Cardillo (2006) who tested segmental contribution to intelligibility in isolated words and found, as did I, results that were inconsistent with Cole et al.

In Owren and Cardillo's study, also on English, consonantal information was found to be primarily responsible for lexical information (word meaning) while vocalic information was responsible for indexical information (the talker's gender). This conclusion about consonantal information concurs with our findings about the role of consonants in speech intelligibility. Consonantal, rather than vocalic, information provided the participants in our study with the necessary information they needed to comprehend speech. The results are also in agreement with those on English from van Ooijen (1996) and Cutler et al. (2000). Although their studies were not a direct investigation of segmental contribution, they both showed that consonantal information is robust and constrains lexical selection more tightly than vocalic information does. Fogerty et al. (2012), however, examined segmental contributions to intelligibility in isolated English words and concluded that the contributions of consonantal and vocalic information appear to be nearly equal at the word level. However, they found that consonants at the word level resulted in better intelligibility than at

the sentence level, while vowels gain *added importance* at the sentence level and their contributions are *context-dependent*.

Their conclusion, however, raises one important question that is relevant to the current study. Fogerty et al. (2012) seem to support the hypothesis that consonants are more important/informative for lexical access and hence for speech intelligibility at the word level, primarily for their robust spectral information that is not subject to context factors. This suggests that consonantal information contributes more to speech intelligibility primarily for the phonetic/acoustic information consonants possess. This could be the case, because English does not distinguish between vowels and consonants in terms of their morphological roles. One may then ask if the results of the current experiment on Arabic are due to the phonetic/acoustic properties of consonants, or due to their morphological properties. The current experiment is not enough to answer this question in its entirety, since the patterns found here are consistent with previous studies that were conducted in concatenative languages. Nevertheless, the results regarding S-Ratio provide some preliminary interpretations that partially support the role of the morphological system. Experiment 2 pursues these questions further.

3.1.3.2 Segmental ratio. The current experiment shows that a word's S-Ratio significantly affects its intelligibility. For isolated words in the CO condition, speech intelligibility was higher in 3C-2V than in in 3C-3V and in 2C-3V. This outcome is interesting for two reasons. First, it shows that the S-Ratio in a word is an important factor in speech intelligibility, which makes us question how much influence it had

in previous studies in which this factor was not controlled for. Second, it highlights the distinction between vocalic information and consonantal information in speech intelligibility. If we take the 3C-3V condition as a baseline, missing one vowel raised performance by 17.50% (= 82.50% intelligibility in 3C-2V condition) whereas missing one consonant lowered performance by 17.50% (= 47.50% in 2C-3V). This pattern of results shows superiority for consonantal information, especially given that a tri-consonantal structure is the most frequent root type in Arabic while bi-consonantal roots are less frequent (see, e.g., Ravid & Shimron, 2003). However, while this may not be taken as a definitive answer to whether morphology plays a role in the speech intelligibility of nonconcatenative languages, the decline of speech intelligibility in the 2C-3V condition can support such a claim. Specifically, this pattern shows that the presence of more vocalic information in the 2C-3V did not help the participants to gain better intelligibility compared to when more consonantal information was present (as in 3C-2V). In other words, 3C-2V and 3V-2C have the same magnitude of information (5 segments each), but the type of segments affected speech intelligibility. The only apparent reason behind this large difference (82.5% - 47.5% = 35%) is an effect of S-Ratio.

Most similar studies in the literature did not explicitly discuss S-Ratio as a factor although their descriptions of stimulus construction can shed some light on this issue. The stimuli in Fogerty et al. (2012) consisted exclusively of CVC words. As such, the information present in the CO and VO conditions was not balanced. We would thus expect speech intelligibility to suffer in the VO condition, not only because

consonantal information contributes more to speech intelligibility but also because S-Ratio played an (overlooked) role. Indeed, when words were placed in the VO condition, 66.66% (two thirds) of the word information was absent while only 33.33% (one third) of word information was absent in the CO condition. Other studies were either unclear about their stimuli's S-Ratios (e.g., Cole et al., 1996; Kewley-Port et al., 2007) or had experimental designs/tasks that did not completely rule out related concerns. For example, Cutler et al. (2000) balanced the number of vowels and consonants in a third of the materials so that the participants had equal opportunity to substitute vowels or consonants. However, this does not eliminate the S-Ratio concern because the task in Cutler et al. involved substituting a segment to shift nonsense words into real words. Other studies that used artificial speech took into consideration S-Ratio (see e.g., Bonatti et al., 2005) but differed in many other theoretical and experimental aspects. For instance, although Bonatti and colleagues used trisyllabic nonsense items such as /gobydo/ and found that consonants were more informative than vowels, they placed items in streams of continuous artificial speech, which makes comparison with isolated real words less than ideal.

Another related source of concern in the current study, as in other studies, is the initial segment in each item. When the initial segment is replaced with silence or noise, the speech loses a critical segment used to determine the set of activated candidates in lexical access (e.g., Marslen-Wilson, 1987). Although I was aware that the first segment being absent could have some effects on intelligibility, it was impossible

to control for this concern in the Arabic stimuli used here, because MSA words generally do not begin with a vowel. As a partial remedy, I placed words between two tones, and explicitly informed the participants that some segments would be absent from speech and that they should attempt to understand the speech to the best of their abilities. Note that we were able to address this issue in English, which will be fully discussed in Experiment 4.

In summary, consonants provide more information for speech intelligibility than vowels do at the word level in Arabic, which stands in agreement with previous findings on concatenative languages. However, we cannot safely attribute their superiority here to the morphological system of Arabic because the results from both concatenative and nonconcatenative languages exhibit the same pattern. Therefore, Experiments 2 and 3 test this hypothesis at the sentence level in Arabic.

3.2 Experiment 2

Experiment 1 examined the speech intelligibility of isolated words in Arabic to investigate the effects of segment type and ratio in nonconcatenative languages. Experiment 2 then addresses the same questions but at the sentence level. It is crucial to distinguish between these two levels because it is at the sentence level where results in Arabic would be expected to diverge from results in other languages. In particular, while vocalic information contributed more than consonantal information to speech intelligibility at the sentence level in languages such as English, Dutch, and Spanish, the reverse was expected to occur in Arabic due to its morphological system. To examine this, I conducted a second experiment in which the same technique from

Experiment 1 was implemented but at the sentence level.

3.2.1 Methodology

3.2.1.1 Overview. In Experiment 2, speech intelligibility was examined at the sentence level. Sentence intelligibility was determined based on the intelligibility of the target words once and based on the intelligibility of the entire sentence once. The stimuli consisted of 48 MSA sentences, and each sentence consisted of exactly 6 words and had a 1:1 S-Ratio – 23 consonants and 23 vowels total, including the target words. The same 24 N vs. V words from Experiment 1 were embedded as target words in 48 the sentences to test one additional factor that was not tested in the literature, which is the word initial vs. final position. Specifically, a sentence provides a meaningful context in which listeners can predict and/or interpret the meaning of a word. This context differs depending upon where the target word occurs within the sentence: early words are typically unpredictable, while later words are typically much more predictable. In order to test for contextual benefits that a word may gain in final position compared to initial position, each target word was placed in two sentences: one in which it occurred sentence-initially, and one in which it occurred sentence-finally. This made up a total of 4 comparisons: CO vs. VO (which is still the major comparison here), N vs. V, sentence-initial vs. sentence-final, and 2C-3V vs. 3C-3V vs. 3C-2V. The same speaker who recorded the stimuli in Experiment 1 recorded the stimuli for Experiment 2, the same silence replacement procedure was also applied here, and the same participants from Experiment 1 participated in Experiment 2, following the

same procedure in Experiment 1. Responses were scored based on correct identification of the target word only, and assigned 1 point if identified correctly and zero points otherwise.

3.2.1.2 Sentence length. The stimuli consisted of 48 sentences (see Appendix I), the exact number of sentences used in Kewley-Port et al. (2007) and slightly more than the number of sentences (42) used in Fogerty, et al. (2012). Each of the 48 sentences in this experiment consisted of exactly six words that represented different parts of speech. This decision was based on Kewley-Port et al. (2007), whose sentences ranged from 6 to 10 words with an average of 8.01 words per sentence. Since Arabic enjoys a rich morphology that allows case, mood, agreement, (in)definite pronouns, and clitics to all be part of what can be counted as one word, a 6-word sentence would likely be close to the average number of English words used in Kewley-Port et al. (2007). It was important that the current design kept the number of words equal across sentences to reduce the diversity in length among the sentences.

3.2.1.3 Sentence construction. As I selected the words used in Experiment 1, I also constructed the sentences used in Experiment 2. I positioned every target word from Experiment 1 in two different sentences, once in sentence-initial position and once in sentence-final position. I also controlled the number of segments preceding or following the target word. Every target word was preceded and followed by exactly the same number of words, and exactly the same number of vowels and consonants. The sentences covered a wide range of topics familiar to an Arabic speaker

such as politics, sports, and life events. All consonants and vowels in Arabic's phonemic inventory were represented not only cross the sentence stimuli, but also cross the target word stimuli (see Appendix J).

The example below demonstrates two sentences from the stimuli. The verb /ʃaarak/ "joined/participated" is placed in sentence-initial position in (1) and in sentence-final position in (2). In both sentences, it is followed or preceded by the same number of consonants and vowels (20 vowels and 20 consonants). The target word itself is of 3C-3V type.

1) **ʃaarak**-a jaħjaa ħamiid-an fii muʔtamar-i l-lisaanijjaat

$c^1v_1v_2c^2v_2c^3\text{-}v_4$ | $c^4v_5c^5c^6v_6v_7$ | $c^7v_7c^8v_9v_{10}c^9v_{11}c^{10}$ | $c^{11}v_{12}v_{13}$ | $c^{12}v_{14}c^{13}c^{14}v_{15}c^{15}v_{16}c^{16}v_{17}$ | $c^{17}c^{18}v_{18}c^{19}v_{19}v_{20}c^{20}v_{21}c^{21}c^{22}v_{22}v_{23}c^{23}$

participate.PRF-3PSM | Yahya-ACC | Hamid-ACC | in | conference-GEN | the-linguistics

"Yahya participated with Hamid in the linguistics conference."

2) lam ja-ktariθi l-fariiq-u bi-l-xasaar-at-i ħiinamaa **ʃaarak**

$c^1v_1c^2$ | $c^3v_2c^4c^5v_3c^6v_4c^7v_5$ | $c^8c^9v_6c^{10}v_7v_8c^{11}v_9$ | $c^{12}v_{10}c^{13}c^{14}v_{11}c^{15}v_{12}v_{13}c^{16}v_{14}c^{17}v_{15}$ | $c^{18}v_{16}v_{17}c^{19}v_{18}c^{20}v_{19}v_{20}$ | $c^{21}v_{21}v_{22}c^{22}v_{23}c^{23}$

not.PRF | care | the-team-NOM | about-the-loss-GEN | when | participate.PRF

"The team did not care about the [potential] loss when it participated"

3.2.1.4 Judgements. I submitted all sentences to 10 native speakers of Arabic to judge them for naturalness and likelihood. I asked the judges to rate, on a 1 to 7 scale, every sentence according to how natural it sounded to them, and again based on how likely it would be heard in their daily lives. The same criterion for inclusion in Experiment 1 was adopted here. Table 11 shows the rating means in both written

and spoken forms (see Appendix B for the entire ratings output). None was judged to have any issue and hence no new recording took place.

Table 11. Judgment ratings in Experiment 1

Form	Sentence Likelihood	Sentence Naturalness	Word Familiarity
Written	6.70	6.91	6.95
Spoken	6.84	6.93	6.96

3.2.1.5 Recording and stimulus preparation. The same steps were followed as in Experiment 1. One additional step was taken to remove pauses from the recording. The speaker of the stimuli made very short pauses between constituents in a few sentences. I decided to remove such pauses before proceeding with silence replacement so that they would not intervene with silence substitution for the segments – the pause and silence replacement will accumulate and produce a long silence. Each sentence was about 6.13 seconds (M = 1.02 word per second), and each sentence articulation rate was about 7.54 phones per second (see Appendix K for the entire duration and articulation rates).

3.2.1.6 Task. The same task and instructions given in Experiment 1 were also adopted in Experiment 2. However, one additional instruction was given in Experiment 2, directing the participants to report whatever they could understand regardless of whether it was a complete and meaningful sentence or not.

3.2.1.7 Participants. The same 20 respondents who participated in Experi-

ment 1 participated in Experiment 2. Half of the participants started with this experiment while the other half started with the previous experiment.

3.2.1.8 Procedure. There was no difference in the procedure between Experiment 1 and Experiment 2 other than the additional instructions as described above.

3.2.1.9 Data scoring and analysis. I scored the data from this experiment twice. First, I used the same scoring technique from Experiment 1. Note that intelligibility of sentences was scored based exclusively on correct recognition of the target word (binary variable: *intelligible* [1] vs. *unintelligible* [0]), irrespective of the rest of the sentence in which the word was contained. Second, I scored the overall intelligibility of the entire sentence (number of intelligible words) regardless of the target words, a scoring technique that similar studies (e.g., Kewley-Port et al. 2007; Fogerty et al., 2012) have used. The total number of intelligible words, which is a type of count data, was divided by the total number of words in each sentence to have an intelligibility percentage for each sentence. Then, each proportion was transferred into angular values and ultimately to RAU, which can be treated as a continuous variable. The first scoring was appropriate for repeated-measures logistic regression and hence for comparison with the results from Experiment 1, and the results for this scoring method were named "target-word results". The second scoring was appropriate for the repeated-measure ANOVA and hence for comparison with results from previous studies as well as from Experiment 3, and the results for this scoring technique were named "entire-sentence results".

3.2.2 Results

Since we have two techniques for scoring, I will break the results into two categories: target-word results and entire-sentence results.

3.2.2.1 Target-word results. As in Experiment 1, for descriptive purposes only, I computed the proportions of intelligibility, converted them to RAU, and presented the outputs once as percentages and once as RAUs. Here I will only emphasize the portions in which differences between conditions exhibit distinct patterns with statistical support.[15] As shown in Table 12, the main results demonstrate higher intelligibility in the CO condition than in the VO condition.

Table 12. Intelligibility (% & RAU) broken down by CO vs. VO in Experiment 2 (Target-Word)

CO vs. VO	Responses	Overall (%)	Overall (RAU)
CO	240	65.11%	64.09
VO	240	45.57	48.09

Comparing the additional factor, Initial vs. Final (IN vs. FI) position, intelligibility proportions were almost equal: 57.45% (56.92 RAU) in IN position vs. 55.92% (55.49 RAU) in FI position (see Appendix L). However, intelligibility in the CO condition was higher for the FI position (69.67%, 68.62 RAU) than for the IN position (60.18%, 59.44 RAU), and lower for the FI position (42.28%, 42.85 RAU) than for the IN position (54.92%, 54.45 RAU) in the VO condition. Similarly, intelligibility in N

[15] For other conditions in which this is not the case, visual illustrations can be found in the appendices (Appendix L).

(55.74%, 55.32 RAU) and V (57.55%, 55.59 RAU) showed a minor difference, but nouns were better identified (68.47%, 67.435 RAU) than verbs (44.35%, 44.78 RAU) in the CO condition, while verbs were better identified (52.89%, 52.66 RAU) than nouns (44.35%, 44.78 RAU) in the VO condition. Nevertheless, the differences in the N vs. V and IN vs. FI conditions were not as large as those in the CO vs. VO. Intelligibility in the S-Ratio conditions showed similar values especially in the VO condition, as shown in Figure 5 below.

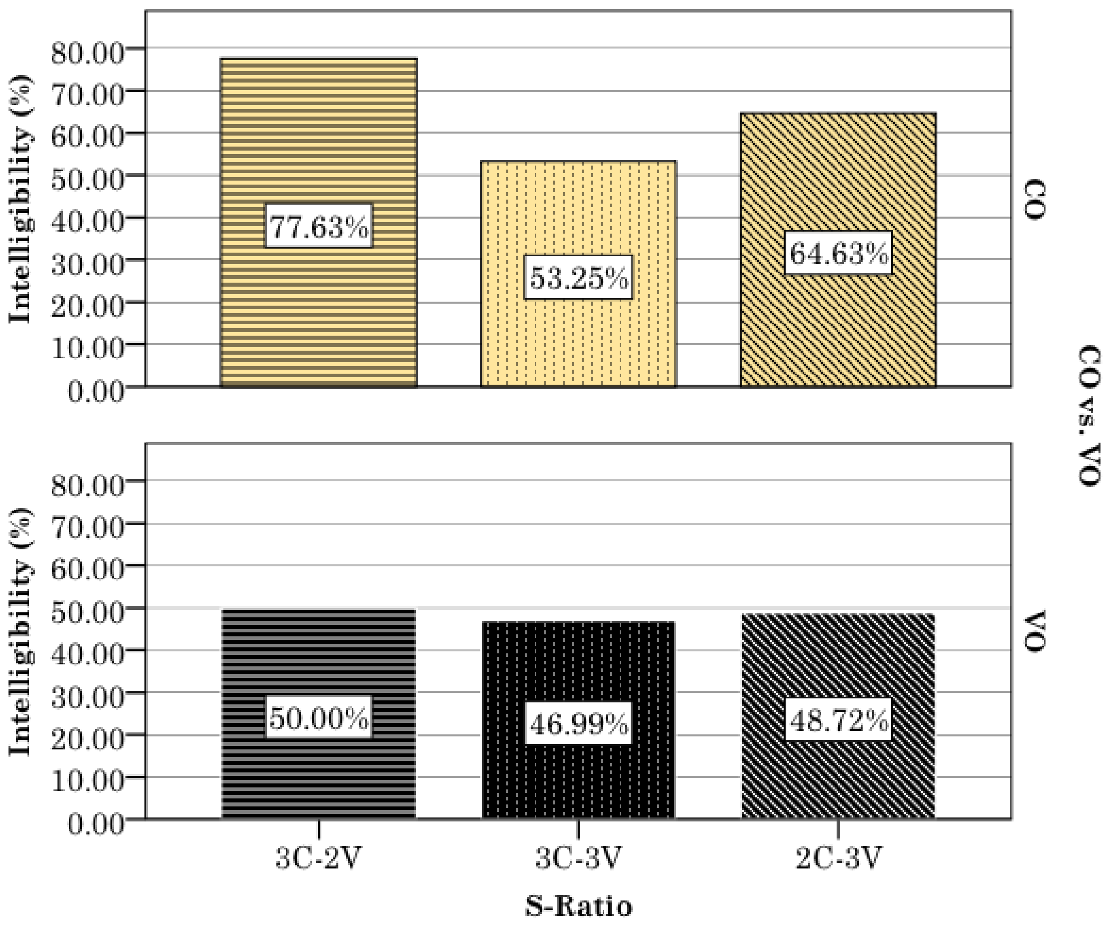


Figure 5. Intelligibility (%) broken down by C0 vs. VO and S-Ratio in Experiment 2

In order to examine the effect of these factors, GEE was used to run a repeated-measures logistic regression as in Experiment 1. The model had subject and trial as repeated effects, CO vs. VO, IN vs. FI, N vs. V, and S-Ratio as predictor variables, and intelligibility as a dependent variable. Other aspects of the model were the same as the model in Experiment 1: Intelligible was the reference category, all 2-way interactions were added to the model, predictor variables were classified as categorical rather than continuous, Wald Chi-square was chosen, Kernel was selected for log quasi-likelihood function, *exchangeable* was selected as a structure for the working correlation matrix (after the goodness of fit table showed that it was the best), and exponential parameter estimates were included (see Appendix M for the model information summary, the correlated data summary, the goodness of fit, and a summary of categorical).

The model produced four main effect and six interaction outputs as main results, while the parameter estimates produced a more detailed output with 9 main effect parameters and 30 interaction parameters. The major results show that CO vs. VO and S-Ratio are two main statistically significant exploratory factors, and that they interact with each other and with other factors, as detailed below. Table 13 provides the summary of the tests of the model effects. The model was able to predict an overall 57.50% of the responses (276 cases) correctly while 42.50% of the responses (201 cases) were misclassified, as shown in the classification table (Appendix M). The model showed that CO vs. VO and S-Ratio statistically significantly predicted whether speech would be in the intelligible vs. unintelligible category, Wald $\chi^2(1)$ =

7.861, p= 0.005, and Wald $\chi^2(2)$ = 14.857, p= 0.001, whereas IN vs. FI and N vs. V did not reach a statistically significant level, Wald $\chi^2(1)$ = 0.001, p= 0.978, and Wald $\chi^2(1)$ = 14.857, p= 0.112, respectively. All two-way interactions were statistically significant, CO vs. VO x IN vs. FI (Wald $\chi^2[1]$ = 7.777, p< 0.005), CO vs. VO x N vs. V (Wald $\chi^2[1]$ = 4.299, p= 0.038), CO vs. VO x S-Ratio (Wald $\chi^2[2]$ = 9.728, p= 0.008), IN vs. FI x N vs. V (Wald $\chi^2[1]$ = 8.390, p= 0.004), IN vs. FI x S-Ratio (Wald $\chi^2[2]$ = 7.337, p= 0.026), and N vs. V x S-Ratio (Wald $\chi^2[2]$ = 26.630, p< 0.000).

Table 13. Summary of the tests of model effects in Experiment 2

Source	Wald χ^2	*df*	Sig.
(Intercept)	1.810	1	.179
CO vs. VO	7.861	1	**.005**
IN vs. FI	.001	1	.978
N vs. V	.112	1	.738
S-Ratio	14.857	2	**.001**
CO vs. VO x IN vs. FI	7.777	1	**.005**
CO vs. VO x N vs. V	4.299	1	.038
CO vs. VO x S-Ratio	9.728	2	**.008**
IN vs. FI x N vs. V	8.390	1	**.004**
IN vs. FI x S-Ratio	7.337	2	**.026**
N vs. V x S-Ratio	26.630	2	**.000**

Dependent variable: Intelligibility

Model: (Intercept), CO vs. VO, IN vs. FI, N vs. V, S-Ratio, CO vs. VO x IN vs. Final, CO vs. VO x N vs. V, CO vs. VO x S-Ratio, IN vs. FI x N vs. V, IN vs. FI x S-Ratio, N vs. V x S-Ratio

We can now examine the individual parameter estimates shown in Table 14. As in Experiment 1, the table has many parameters (nine main effect parameters and 30 interaction parameters). Here only the significant results will be discussed (for more see Appendix N). Table 15 shows *unintelligible* as the reference category for easier interpretations when b<0 and/or Exp(B)<1. As Table 15 shows, speech in the VO category was predicted to be in the unintelligible category 4.422 times (≈ 81.55%) higher than speech in the CO category, b= 1.487, SE= 0.5528, 95% CI = [0.403, 2.570], Wald $\chi^2(1)$ = 7.231, p= 0.007, Exp(B)= 4.422 (≈ 81.55%), 95% CI = [0.077, 0.668]. It is therefore (more than four times) more likely for VO speech to be in the unintelligible category compared to CO speech. There are five statistically significant interaction parameters: CO x FI, b= .914, SE= 0.3278, 95% CI = [0.272, 1.557], Wald $\chi^2(1)$ = 7.777, p= 0.05, Exp(B)= 2.495 (≈ 71.83%), 95% CI = [1.312, 4.744]; CO x V, b= 0.532, SE= 0.2565, 95% CI = [0.029, 1.035], Wald $\chi^2(1)$ = 4.229, p= 0.038, Exp(B)= 1.702 (≈ 62.99%), 95% CI = [1.030, 2.814]; FI x V, b= 1.161, SE= 0.4008, 95% CI = [0.375, 1.946], Wald $\chi^2(1)$ = 1, p= 0.004, Exp(B)= 3.193 (≈ 76.15%), 95% CI = [1.456, 7.003]; FI x 3C-2V, b= 1.007, SE= 0.3752, 95% CI = [0.272, 1.742], Wald $\chi^2(1)$ = 7.203, p= 0.007, Exp(B)= 2.737 (≈ 73.24%), 95% CI = [1.312, 5.710]; and V x 3C-3V, b= 1.732, SE= 0.3379, 95% CI = [1.070, 2.394], Wald $\chi^2(1)$ = 26.279, p= 0.001, Exp(B)= 5.652 (≈ 84.96%), 95% CI = [2.915, 10.959].

In summary, CO vs. VO and S-Ratio were consistently explanatory factors and interacted with different levels of other factors in the model effects, predicting response membership to the intelligible and unintelligible categories. Speech from the

CO category was more likely to be in the intelligible category than that from the VO category. On the other hand, the N vs. V and IN vs. FI factors did not show statistically significant effects on the model, although they showed some interaction effects with each other and with different levels of other factors. Since these two factors were not overall predictors as seen in the tests of the model effects, and since the S-Ratio, N vs. V, and IN vs. FI were characteristics of the target-word rather than the entire sentence, I decided to run another test to examine the overall intelligibility for the sentences in both the CO and VO conditions. In doing so, I scored the intelligibility of the entire sentence, transformed it into angular values and then into RAUs, and then ran some statistics, as detailed in 3.2.2.2. This made the results more comparable to those in Experiment 1 and in the literature. In other words, scoring sentential intelligibility was necessary in order to compare it with word intelligibility, which is one of the major comparisons in this study.

Table 14. Summary of the GEE parameters in Experiment 2 (Ref. Category= Intelligible)

Parameter	B	Std. Error	95% Wald CI		Hypothesis Test			Exp(B)	95% Wald CI for Exp(B)	
			Lower	Upper	Wald χ^2	*df*	Sig.		Lower	Upper
(Intercept)	.218	.4163	-.598	1.033	.273	1	.601	1.243	.550	2.811
[CO vs. VO=2.00]	-1.487	.5528	-2.570	-.403	7.231	1	**.007**	.226	.077	.668
[CO vs. VO=1.00] x [IN vs. FI=2.00]	.914	.3278	.272	1.557	7.777	1	**.005**	2.495	1.312	4.744
[CO vs. VO=1.00] x [N vs. V=2.00]	.532	.2565	.029	1.035	4.299	1	**.038**	1.702	1.030	2.814
[IN vs. FI=2.00] x [N vs. V=2.00]	-1.161	.4008	-1.946	-.375	8.390	1	**.004**	.313	.143	.687
[IN vs. FI=2.00] x [S-Ratio=1.00]	-1.007	.3752	-1.742	-.272	7.203	1	**.007**	.365	.175	.762
[N vs. V=2.00] x [S-Ratio=2.00]	1.732	.3379	1.070	2.394	26.279	1	**.000**	5.652	2.915	10.959

Dependent variable: Target-Word Intelligibility

Reference category: *Intelligible*

Model: (Intercept), CO vs. VO, IN vs. FI, N vs. V, S-Ratio, CO vs. VO x IN vs. FI, CO vs. VO x N vs. V, CO vs. VO x S-Ratio, IN vs. FI x N vs. V, IN vs. FI x S-Ratio, N vs. V x S-Ratio

Table 15. Summary of the GEE parameters in Experiment 2 (Ref. Category= Unintelligible)

Parameter	B	Std. Error	95% Wald CI		Hypothesis Test			Exp(B)	95% Wald CI for Exp(B)	
			Lower	Upper	Wald χ^2	*df*	Sig.		Lower	Upper
(Intercept)	-.218	.4163	-1.033	.598	.273	1	.601	.804	.356	1.819
[CO vs. VO=2.00]	1.487	.5528	.403	2.570	7.231	1	**.007**	4.422	1.496	13.068
[CO vs. VO=1.00] x [IN vs. FI=2.00]	-.914	.3278	-1.557	-.272	7.777	1	**.005**	.401	.211	.762
[CO vs. VO=1.00] x [N vs. V=2.00]	-.532	.2565	-1.035	-.029	4.299	1	**.038**	.587	.355	.971
[IN vs. FI=2.00] x [N vs. V=2.00]	1.161	.4008	.375	1.946	8.390	1	**.004**	3.193	1.456	7.003
[IN vs. FI=2.00] x [S-Ratio=1.00]	1.007	.3752	.272	1.742	7.203	1	**.007**	2.737	1.312	5.710
[N vs. V=2.00] x [S-Ratio=2.00]	-1.732	.3379	-2.394	-1.070	26.279	1	**.000**	.177	.091	.343

Dependent variable: Target-Word Intelligibility

Reference category: *Unintelligible*

Model: (Intercept), CO vs. VO, IN vs. FI, N vs. V, S-Ratio, CO vs. VO x IN vs. FI, CO vs. VO x N vs. V, CO vs. VO x S-Ratio, IN vs. FI x N vs. V, IN vs. FI x S-Ratio, N vs. V x S-Ratio

3.2.2.2 Entire-sentence results. As shown in Table 16 (see Appendix O for visual illustrations), the overall results show that speech intelligibility was higher in the CO condition (75.06% and 77.60 in RAU) than in the VO condition (53.11% and 52.58 in RAU).

Table 16. Intelligibility (% & RAU) broken down by CO vs. VO in Experiment 2 (Entire-Sentence)

CO vs. VO	Responses	Overall (%)	Overall (RAU)
CO	240	75.06%	77.06
VO	240	53.11%	52.58

Figure 6 shows the mean, median, range, and distribution of intelligible words. The figure shows that the medians in the CO and VO conditions are six and three, respectively, and that in most cases at least one word from the CO condition was intelligible, whereas speech from the VO condition was sometimes completely unintelligible.

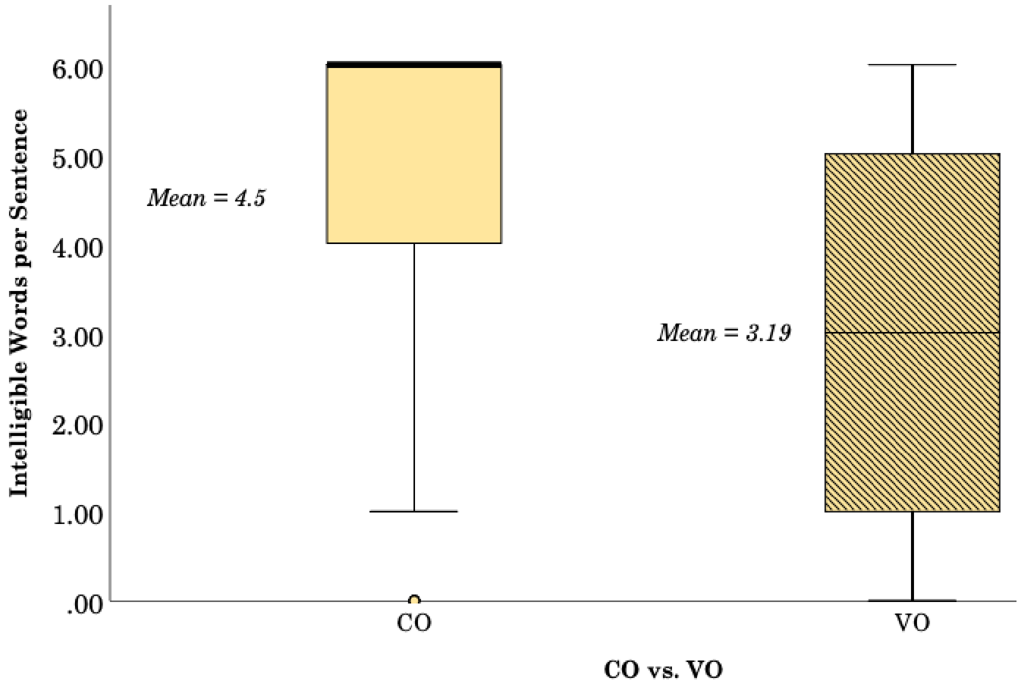


Figure 6. Intelligibility (# of words) broken down by C0 vs. VO in *Experiment 2*

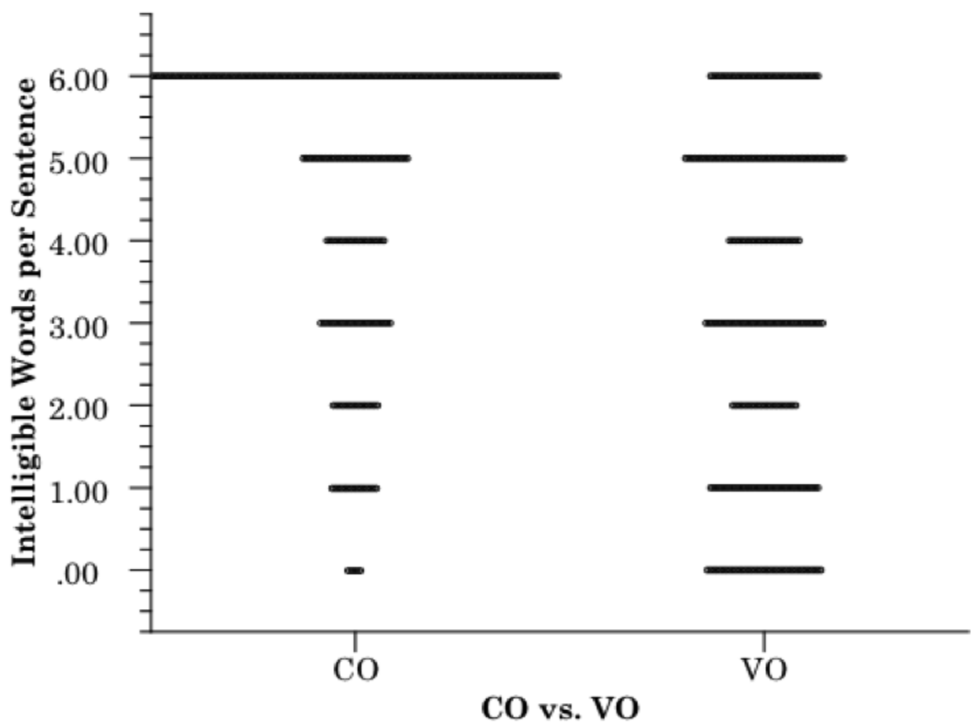


Figure 7. Dot-plot illustration for the frequencies of words per sentence in *Experiment 2*

Another way to investigate the frequencies and distributions is to scrutinize the percentages of intelligibility (see Appendix P for frequencies presented on histograms). Although it was noticed that a few sentences were more intelligible than all other sentences in both the CO and VO conditions, the sentences in the CO condition were 100% intelligible to the participants 129 times compared to only 35 times in the VO condition. The sentences were completely unintelligible (00.00%) only 5 times in the CO condition and 37 times in the VO condition. Intelligibility percentages were skewed to the left in the CO condition, but nearly symmetric in the VO condition. This raises some challenges for using the mean in statistical tests, as the data shows a degree of non-normal distribution. When data from both conditions were grouped as one category, there was still a generally left-skewed distribution, implying that the data from the CO condition is dominating the entire data distribution. This raises some questions as to which would be the most appropriate statistical test, since the normality assumption may not be met. To deal with this concern, and as detailed below, I ran a normality test, computed the skewedness and kurtosis of the residuals for both CO and VO RAUs, and decided that the repeated-measures ANOVA was nonetheless well-suited for examining the differences between the two conditions under investigation.

Despite this skewed normality distribution, with only one dependent variable ANOVA is usually robust and tolerates assumption violations quite well. There are typically two solutions to deal with normality concerns. First, data can be trans-

formed into different values to reduce skewedness or kurtosis. The data was originally countable (number of words), and the dot-plot chart showed that at least the data in the CO condition were non-normal. The data were transformed into percentages, but still showed a non-normal shape. Ultimately, the percentages were transformed into RAU. Second, normality can simply be assumed, and further tests of normality may be performed if the results show statistically significant differences. In the current study, a few steps were taken to examine normality before running any statistical test. The first step was to run a Shapiro-Wilk normality test, the outputs of which are given below.

The Shapiro-Wilk (as well as Kolmogorov-Smirnov) test showed that there was evidence of a statistically significant difference between the current RAU data distribution and normality distribution, p= 001. We could use the proportional values in lieu of RAUs to check for normality and run ANOVA, but it is unlikely we would obtain different results. Therefore, there is no need to test normality using percentages or even number of words. However, considering the nature of the data the normality test is unlikely to be determinant in this case. The original data were countable and only seven outcomes were possible (from 0 to 6). Transformation into percentages added some decimal values and expanded the range of data. Transformation into RAUs then added negative values, but the seven values were repeated multiple times in the results. Therefore, normality tests are very sensitive to repeatable values, which leaves the question about normality yet unanswered.

I decided instead to compute values for skewedness and kurtosis of the residuals for both CO and VO RAUs (see Appendix Q for more normality tests and illustrations), since normality tests may fail to reject the null hypothesis due to the data's nature (the data are count-like) and sample size (large samples are required). The data from the CO condition have more skewedness (-1.194) but less kurtosis (0.0491) while the data from the VO condition show the opposite (-0.155 vs. -1.03, respectively). There are no concerns about the two values, 0.0491 and -0.155, because both range between -0.5 and 0.5. The two other values, -1.194 and -1.03, are still acceptable because they are close to -/+1 and neither reached -/+2 (see Appendix Q).

In summary, although the distribution cannot be deemed normal, there are several reasons to support running such a test regardless of normality. First, with one dependent factor, ANOVA can overcome normality violations. Second, normality tests of the residuals did not produce large values. Third, the results will be comparable to those from previous studies using means to run ANOVA. However, it is wise to run another test to avoid running into the risk of obtaining an uncertain answer to one of the main research questions due to statistical complexity. Although the RAUs appear as a continuous variable, one cannot neglect the fact that they were transformed from count data. Additionally, the same normality dilemma may engender accumulating uncertainty in the following experiments, which may require running multiple tests each time. Therefore, I decided to run two tests as detailed below.

I first performed a General Linear Model repeated-measures ANOVA test. The

independent variable was CO vs. VO while the dependent variable was RAU, although similar studies (e.g., Kewley-Port et al., 2007) used the total number of words per sentence to run a repeated-measures ANOVA. The ANOVA test indicated that there was a significant main effect for the factor under investigation at the 0.001 level, $F(1, 19) = 38.311$, $p= 0.001$, $\omega^2 = 0.66$. The effect size was large as shown by the partial eta squared in the last column in Table 17, $\omega^2 = 0.66 > 0.14$, which in fact equals the eta squared value because there was only one independent factor. The table also shows different corrections such as Greenhouse-Geisser, although sphericity was automatically met because there were only two levels of the independent factor.

Table 17. Summary of repeated-measures ANOVA in Experiment 2 (entire sentence)

Source		Type III Sum of Squares	*df*	Mean Square	F	Sig.	Partial Eta Squared
CO-VO	Sphericity Assumed	92304.308	1	92304.308	38.311	**.000**	.66
	Greenhouse-Geisser	92304.308	1.000	92304.308	38.311	**.000**	.66
	Huynh-Feldt	92304.308	1.000	92304.308	38.311	**.000**	.66
	Lower-bound	92304.308	1.000	92304.308	38.311	**.000**	.66
Error (COVO)	Sphericity Assumed	45777.675	19	2409.351			
	Greenhouse-Geisser	45777.675	19.000	2409.351			
	Huynh-Feldt	45777.675	19.000	2409.351			
	Lower-bound	45777.675	19.000	2409.351			

In addition, GEE (see Appendix R for model details) were used to run a repeated-measures regression analysis using the row number of words as the count dependent variable. The RAUs were not acceptable as a dependent variable for the GEE because tests with Poisson distribution are only used for positive integers. The proportional values could have been used, but there was no convincing reason to use them instead of the original count data – number of words. The candidate test for count data would be a Poisson regression, but we have already seen some forms of non-normal distribution. Instead, I used a negative binomial regression, which is a generalization of Poisson but with an additional parameter to handle any inequalities between means and variances (e.g., overdispersion and underdispersion). CO vs. VO was the independent variable, intelligibility (shown by the number of words) was the dependent variable, subject was the subject variable, trial was the repeated effect, and correlation matrix was set up to *exchangeable*. The regression parameter was left for the model to estimate using the log link function in order to determine the best distribution, which came out as the maximum-likelihood estimation (MLE).

Table 18. Summary of the tests of model effects in Experiment 2

Source	Wald χ^2	*df*	Sig.
(Intercept)	1239.104	1	**0.001**
CO vs. VO	18.448	1	**0.001**

Dependent variable: Intelligibility (number of words)

Model: (Intercept), CO vs. VO

The model showed that CO vs. VO statistically significantly predicted intelligibility, Wald $\chi^2(1) = 18.448$, $p= 0.001$ (see Table 18 above). This output is sufficient to show that CO vs. VO is a statistically significant factor, as we do not have any other factors in the model. Nevertheless, we can still look at the effect of speech in the CO vs. VO conditions on the dependent variable. The individual parameters table (Table 19) shows that speech in the CO category was significantly predicted to have a *log* count 1.507 times (≈ 60.11%) higher than in the VO category, $b= 0.410$, $SE=$ 0.0955, 95% CI = [0.223, 0.597], Wald $\chi^2(1) = 18.448$, $p= 0.001$, Exp(B)= 1.507 (≈ 61.11%), 95% CI = [1.250, 1.818].

In summary, both types of statistical tests, regression and ANOVA, showed that CO vs. VO is a predictor of speech intelligibility. S-Ratio of target words was shown by the regression model as secondary factor in speech intelligibility. All other factors did not significantly contribute to the model, but individual parameters showed that such factors had some main and interaction effects with each and with both the CO vs. VO and S-Ratio.

Table 19. Summary of the GEE parameters in Experiment 2 (Entire-Sentence)

Parameter	B	Std. Error	95% Wald CI		Hypothesis Test			Exp(B)	95% Wald CI for Exp(B)	
			Lower	Upper	Wald χ^2	*df*	Sig.		Lower	Upper
(Intercept)	1.136	.0719	.995	1.277	249.918	1	**.000**	3.114	2.705	3.586
[CO vs. VO=1.00]	.410	.0955	.223	.597	18.448	1	**.000**	1.507	1.250	1.818
[CO vs. VO=0.00]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)										
(Negative Binomial)	4.152E-8b									

Dependent variable: Entire-Sentence Intelligibility

Model: (Intercept), CO vs. VO

3.2.3 Discussion and Conclusion

Experiment 2 was specifically designed to examine segmental contributions and the effects of S-Ratio on speech intelligibility at the sentence level in Arabic. The overarching results from both analyses (target-word and entire-sentence) can be summarized in a few points. First, CO vs. VO was a strong and persistent predictor of speech intelligibility. Second, although each sentence had equal numbers of consonants and vowels, this balanced S-Ratio at the sentence level did not manifest itself in a balanced way with respect to speech intelligibility in the CO and VO conditions. Rather, intelligibility was higher when participants were presented with only consonantal information than when presented with only vocalic information. Third, considering the word level only, compared to the results in Experiment 1, different S-Ratios did not trigger large differences in speech intelligibility as shown by both percentages and RAUs. Fourth, neither the position of the target word in the sentence nor its part of speech (noun vs. verb) was an explanatory factor for speech intelligibility. These findings are divergent from those in previous studies about concatenative languages, shedding light on segmental contributions unique to nonconcatenative languages as outlined in the subsequent sub-sections.

3.2.3.1 Segmental contributions. As can be seen, the differences between our results and those of previous studies are dramatic. For example, Cole et al. reported zero 100% intelligibility when only obstruents or weak sonorants (i.e., consonants) were preserved in speech for sentences. In a similar vein, Fogerty et al. (2012) re-

ported approximately 18.00 (RAU) intelligibility when only consonants were preserved, and Kewley-Port et al. (2007) concluded that consonantal information is only responsible for approximately 30% of speech intelligibility in non-noisy environments. In the current experiment, sentences were wholly intelligible (100%) 129 times in the CO condition and entirely unintelligible (00.00%) only 5 times in the VO condition.

It is not surprising that the pattern here is different from that reported in the literature given the unique morphological system of Arabic. All previous studies used stimuli from languages that do not assign different morphological roles to consonants vs. vowels, such as English, Dutch, and Spanish. Thus, these studies were obliged to focus primarily on the acoustic information conveyed by consonants vs. vowels, and not their morphological properties. Hence, it seems reasonable that their conclusions about segmental contributions were overwhelmingly phonetics-based. Arabic, as well as other languages with nonconcatenative systems, clusters its lexicon around the consonantal root used to organize, access, and recognize lexical items (see e.g., Abu-Rabia, 1997; Bentin & Feldman 1990; Deutsch, Frost, Pollatsek, & Rayner, 2000; Deutsch & Frost, 2003; Feldman, Frost, & Pnini, 1995; Holes, 2004; Minouni et al.1998; and Ravid & Shimron, 2003 among many others). This suggests a "consonantal privilege" that allows listeners to recognize and process words while relying primarily on the information available in the root. The current experiment, with its high intelligibility in the CO condition, provides further support for this notion. This finding contributes two new insights to our understanding of speech intelligibility.

First, the phonetics-only approach to speech intelligibility is challenged here. The magnitude and type of acoustic information carried by consonants vs. vowels may not alone be responsible for speech intelligibility; rather, the way a language structurally employs each segment type – such as the semantic information in the Arabic root (consonants) vs. the morphosyntactic information in the Arabic pattern (vowels) – is also crucial. This means that morphology intersects with acoustic cues to modulate the contribution of each segment to speech intelligibility.

Second, although the role of the root in mental representations and lexical access has long been recognized by researchers, many previous studies focus on the intelligibility of isolated words rather than entire sentences. That is, evidence for the root as a morphemic unit has been drawn from many morphological experiments using partial or complete words. In the results here, consonantal information as present in the root turned out to be crucial not only at the word level but also at the sentence level. In contrast, the morphological pattern represented by vowels was not sufficient for speech intelligibility at the word level, nor was it as useful as the root for speech intelligibility at the sentence level. This is entirely plausible because a given root can generate only a limited set of possible words, but a given pattern can be applied to a very large number of words. For instance, the root "ʔʃq" from Table 1 can generate different words that are related to LOVE, but the pattern CaaCiC, which is used to form the active participle/doer of something, can be used to form thousands of words so long as the action can have a doer. Thus, it is possible that participants could use

a given pattern to falsely recognize an unintended word. For instance, in the VO condition, some participants provided the response /tˤaaʔir-at-un kabiir-at-un/ "big plane" for the intended two words /ðˤaahir-at-un xatiir-at-un/ "dangerous phenomenon." In contrast, recognition of information from the consonantal root in nonconcatenative languages can help listeners access the target word, while context, position within the sentence, intonation, and other non-phonetic information can assist in determining relevant structural information (e.g., noun vs. verb, singular vs. plural, passive vs. active, etc.) which is carried primarily by vowels. This difference between the role of consonantal information and that of vocalic information in nonconcatenative languages cannot be attributed to an asymmetry in the number of segments in speech, as the S-Ratio was balanced in all sentences. This is thus another aspect of the results that deserves its own discussion.

3.2.3.2 Segmental ratio. S-Ratio of target words at the sentence level was the only statistically significant *overall* predictor of intelligibility in addition to the primary factor, CO vs. VO. In addition, it interacted with CO vs. VO as well as with other factors in the main model. However, looking at individual parameters, none of the S-Ratio levels appeared to have statistically significant main effects or interactions with either the CO or the VO levels. This means that, unlike in Experiment 1, when we look at a level of S-Ratio (e.g., 3C-2V) as an individual parameter or a level of CO vs. VO (e.g., CO) as an interacting parameter, there is statistically not enough evidence to predict intelligibility. In other words, S-Ratio of individual target words embedded in sentences was not as strong as it was for the isolated words in Experiment

1. This can be visually observed in Figure 5 which displays the minor differences between the three levels of S-Ratio in the VO condition. This indicates that vowels at the sentence level provide similar structural information regardless of the target word's S-Ratio, as the entire sentence's S-Ratio was strictly balanced (23C-23V). The results also show slightly larger (but not substantial) differences between the levels of S-Ratio in the CO condition, however the effects of consonants on intelligibility were less consistent than in Experiment 1. This could be a result of the unfixed lexical boundaries in a sentence environment. It is difficult for participants to formulate word boundaries and know which consonant belongs to which root in sentences with half of the segments replaced by silence. Nevertheless, we may still understand the differences between S-Ratio levels in the CO condition as an indication of the differences between the role of consonantal information and that of vocalic information, even if the results did not reach statistical significance. That is, with the same number of vowels and consonants, similarity between the effects of S-Ratio levels in the VO condition would be expected to manifest in the CO condition as well, but this did not occur. Rather, the moderate disparity between the intelligibility rates for the S-Ratio levels in the CO condition may signal a difference between consonants and vowels that were balanced at the sentence level even with an absence of statistical support.

The role of the target word's S-Ratio in intelligibility is impossible to investigate when scoring the intelligibility of entire sentences because each single word con-

tributes some proportional intelligibility to the complete sentence. However, if consonantal information and vocalic information play identical roles, the complete sentence's S-Ratio (balanced: 23C-23V) should have balanced intelligibility rates in the CO condition vs. VO Condition. However, our results show that, even with this balanced S-Ratio, consonantal information still played a greater role in speech intelligibility. This pattern conflicts with the findings from Cole et al. (1996) who claimed that vowels and consonants were balanced in their stimuli.

However, as explained in Chapter 2, Cole et al. (1996) did not really have a balanced S-Ratio partially because they added weak sonorants to the vowel class and partially because they only assessed the total number of vowels and consonants in the whole set of stimuli. Notwithstanding, vowels (be they less than or equal to consonants in their study) played a greater role in speech intelligibility in Cole et al., unlike in our experiment here. To emphasize this, a sentence with fewer vowels in a concatenative language such as English would be more intelligible in the VO condition than in the CO condition compared to a sentence with a balanced S-Ratio in a nonconcatenative language such as Arabic in the VO condition. This leaves us with no clear interpretation other than that consonants and vowels provide different lexical information in concatenative vs. nonconcatenative languages. Consonants provide lexical information and hence play a greater role in intelligibility in concatenative languages due to their morphological systems that favor consonants as a semantic information carrier.

Experiment 2 thus lead to an obvious conclusion that was substantially different from those in previous studies: consonants in complete sentences are more informative to listeners for speech intelligibility in Arabic. The balanced S-Ratio at the sentence level did not equalize speech intelligibility across the CO and VO conditions; rather, speech intelligibility was considerably higher in the CO condition. The next experiment therefore manipulates S-Ratio at the sentence level to examine its impact on speech intelligibility and test whether consonantal information will remain superior in terms of its influence on speech intelligibility in Arabic.

3.3 Experiment 3

If consonants contribute more than vowels to speech intelligibility in nonconcatenative languages, speech intelligibility will suffer in VO conditions where the lexical information carried by consonants is not available to the listener. This is supported by the differences among responses to stimuli in Experiment 1 which had different S-Ratios at the word level. However, Experiment 2 did not vary S-Ratios at the sentence level. In order to complement the two previous experiments and further support the claim that consonants contribute more to speech intelligibility in Arabic than in concatenative languages, Experiment 3 uses a different S-Ratio in complete sentences than Experiment 2 did: namely, in Experiment 3, the S-Ratio of all sentences is set uniformly at 13:23 (0.57:1, vowels to consonants). The results will be compared to those from Experiment 2, in which all sentences had a S-Ratio of 1:1. If the mor-

phological system of Arabic is truly special, Experiment 3's ratio of 13:23, with consonants exceeding vowels by a 10-segment difference, should have a large impact on speech intelligibility compared to Experiment 2's ratio of 1:1.

3.3.1 Methodology

3.3.1.1 Overview. Following the procedures in Experiment 2 as closely as possible, I constructed 24 Arabic sentences, each with 13 vowels and 23 consonants (a 0.57:1 S-Ratio). The same Arabic native speaker produced the sentences, and the same silence replacement was applied to generate two versions, VO and CO, as detailed below.

3.3.1.2 Sentence length. The stimuli consisted of 24 sentences (see Example 3a below) to make the total number of sentences that each subject hears comparable to that of the previous experiment. The average number of words in each sentence was kept approximately the same, *M*= 5.87 words (see Appendix S), but the number of segments differed from that in Experiment 2, as explained below.

3.3.1.3 Sentence construction. Similar to Experiment 2, I constructed the sentences to control for factors such as length and S-Ratio. The number of consonants was identical to that in the previous experiment at 23, while the number of vowels was reduced to 13, for the reason explained above. An ideal design would have sentences consisting of exclusively one type of segment, either vowels or consonants, to reliably reveal the effect of S-Ratio. This is of course impossible, as all languages use both vowels and consonants and Arabic is not an exception. However, it is possible in Arabic to construct sentences with fewer vowels and more consonants, as Example 3

shows below (see Appendix I for all sentences). While the opposite (number of vowels exceeding number of consonants) could be achieved when a sentence incorporates many long vowels, this would affect the naturalness of most sentences and the segmental difference would remain at a minimum (only a few segments) to the extent that ultimately it may not have any impact. Nevertheless, sentences in which the number of consonants exceed the number of vowels should still suffice to examine the hypothesis here, partially because of the opposition between the VO and CO conditions and partially by comparison with the results from Experiment 2. Compare the difference in S-Ratio between Example 3a from the current experiment (Experiment 3) with Example 3b from Experiment 2.

3a) lam | ja-kun | baʕdˁ-u | sˁ-sˁaħb-i | muktariθ-an | bi-l-ʔamr

$c^1v_1c^2$ | $c^3v_2c^4v_3c^5$ | $c^6v_4c^7c^8v_5$ | $c^9c^{10}v_6c^{11}c^{12}v_7$ | $c^{13}v_8c^{14}c^{15}v_9c^{16}v_{10}c^{17}v_{11}c^{18}$ | $c^{19}v_{12}c^{20}c^{21}\mathbf{v_{13}}c^{22}\mathbf{c^{23}}$

not.PRF | 3PSM-be | some-NOM | the-companions-GEN | concerned-ACC | with-the-matter

"Some of the companions were not concerned about the matter"

3b) lam | ja-ktariθi | l-fariiq-u | bi-l-xasaar-at-i | ħiinamaa | **ʃaarak**

$c^1v_1c^2$ | $c^3v_2c^4c^5v_3c^6v_4c^7v_5$ | $c^8c^9v_6c^{10}v_7v_8c^{11}v_9$ | $c^{12}v_{10}c^{13}c^{14}v_{11}c^{15}v_{12}v_{13}c^{16}v_{14}c^{17}v_{15}$ | $c^{18}v_{16}v_{17}c^{19}v_{18}c^{20}v_{19}v_{20}$ | $c^{21}v_{21}v_{22}c^{22}\mathbf{v_{23}c^{23}}$

not.PRF | care | the-team-NOM | about-the-loss-GEN | when | participate.PRF

"The team did not care about the [potential] loss when it participated"

All other construction criteria and decisions regarding experimental design were followed as in Experiment 2.

3.3.1.4 Judgements. I submitted all sentences in written and spoken forms

to the same 10 native speakers of Arabic from Experiment 1 to judge sentences for naturalness and likelihood, using the same criteria as in Experiment 2. Sentences passed the criterion with an overall rate of 6.35 in written form and 6.51 in spoken form (see Appendix T for individual rates).

3.3.1.5 Recording and stimuli preparation. I followed precisely the same steps taken in Experiment 2. The speaker of the new stimuli was the same speaker who recorded the stimuli in Experiments 1 and 2.

3.3.1.6 Task. The participants performed the same tasks performed in the previous two experiments.

3.3.1.7 Participants. Twenty native Arabic speakers participated in the study. The participants come from different Arab countries and speak different dialects. Some of them have participated in the previous experiments, but participants were as close as possible to the continuing participants in terms of age (age ranges from 23 to 42, $M = 33$), the foreign languages they speak, and education level. They had no exposure to the stimuli during any phase of the sentence construction, sentence rating, or stimuli recording.

3.3.1.8 Procedure. I followed the same procedure as in Experiment 2.

3.3.1.9 Data scoring and analysis. In the previous experiment, sentences were scored based on correct identification of the target word and were assigned 1 point if the target word was correctly identified and zero otherwise. In the current experiment, because S-Ratio over complete sentences is the issue under investigation,

I scored responses according to intelligibility of the entire sentence. The scoring procedure therefore was different. Following Fogerty et al. (2012), I assigned sentences a percentage score based on the number of words correctly identified. All word percent-correct scores were then transformed to RAUs and were examined statistically. For comparison, the same scoring steps were also applied to the data from Experiment 2.

3.3.2 Results.

The overall results show that speech intelligibility was higher in the CO condition (90.41% and 91.47 in RAU) than in the VO condition (28% and 27.77 in RAU). Table 20 shows the mean in both conditions (see Appendix U for more illustrations).

Table 20. Overall intelligibility (% & RAU) broken down by CO vs. VO in Experiment 3

CO vs. VO	Responses	Overall (%)	Overall (RAU)
CO	240	91.45	91.47
VO	240	28.00	27.77

As shown in Figure 8, the number of intelligible words per sentence is 5.34 words in the CO condition and 1.65 words in the VO condition. The boxplot (Figure 8) and dot-plot (Figure 9) show almost opposite patterns for the two conditions. The boxplot shows that the number of intelligible words in the CO condition range between four and six words (as represented by the lower and upper whiskers) and 6 is the median. In contrast, it shows that the number of intelligible words in the VO condition ranges between zero and three words and that 1 is the median. The figure also shows more variability in CO compared to VO. The dot-plot shows a similar pattern, but includes the frequency of words in each condition, which, again, shows two opposite patterns. The figure also shows that, although responses with 4 to 6 words were more frequent in the CO condition, there were some responses with 0 to 3 words that were regarded as outliers. More importantly, the figure shows a zero-inflation

pattern in the VO condition, and a heavy tail in the CO condition, hence overdispersion in both conditions. This distribution once again merited a discussion about normality; therefore some further investigation of the data was necessary.

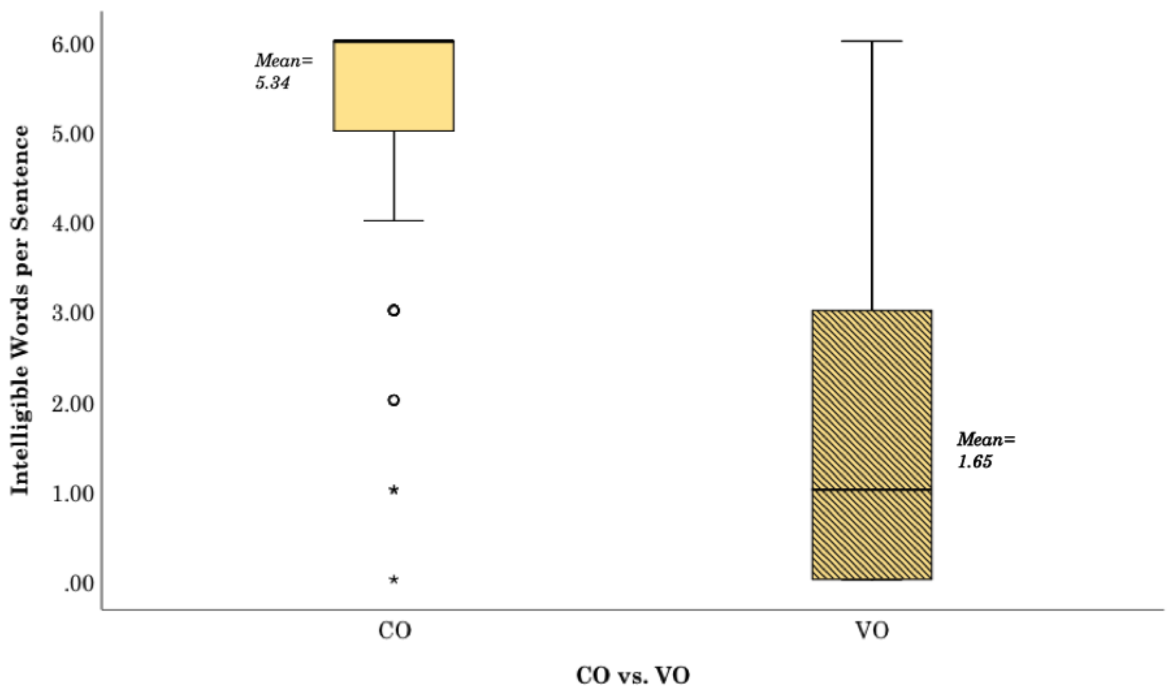


Figure 8. Overall intelligibility (# of words) broken down by CO vs. VO in Experiment 3

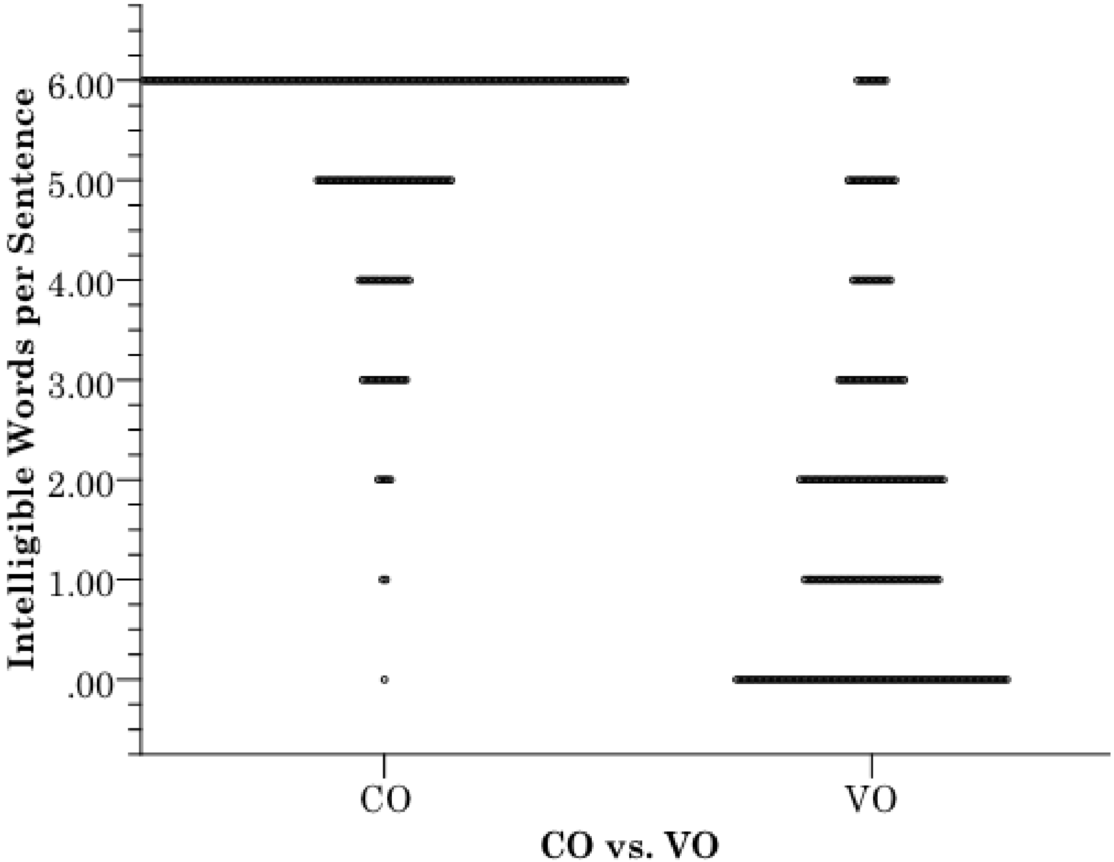


Figure 9. Dot-plot illustration for the frequencies of words per sentence in Experiment 3

Examining the frequencies of percentages (see Appendix V for more details and visual illustrations), sentences were 100% intelligible 170 times in the CO condition, compared to only 10 times in the VO condition. Sentences were also completely unintelligible (00.00%) 88 times in the VO condition. The lowest intelligibility percentage in the CO condition was 17%. Intelligibility percentages are skewed to the left in the CO condition, skewed to the right in the VO condition, and generally U-shaped when data from both conditions are grouped. This is very similar to the situation in Experiment 2, but one difference here is that when we transformed the data from words per sentence to RAUs we obtained more than seven possible outputs. This is because

every sentence in Experiment 2 was 6 words, but in Experiment 3 there were a few stimuli sentences with 5 or 7 words which added more possible percentages. For example, when a participant recognized only 1 out of 6 words in Experiment 2, that participant recognized 16.66% of the sentence, but when a participant recognized 1 out of 5 or 1 out of 7 words in Experiment 3, the participant recognized 20.00% or 14.28% of the sentence, respectively. Therefore, the number of possible percentages multiplied and turned the count data into more continuous data, compared to that in Experiment 2. Taken together, the nature of data became a less serious problem, but the distribution was still an outstanding issue.

As in Experiment 2, a normality test was run in Experiment 3. The Kolmogorov-Smirnov and Shapiro-Wilk tests (see Appendix W) show that there is a statistically significant difference between normally distributed data and the data in Experiment 3, p= 001. I also calculated the values of skewedness and kurtosis and reported the results in (see Appendix W). Data in both conditions have some degree of kurtosis and skewedness but are more platykurtic in the CO condition (-1.548) than in the VO condition (-0.442), and more skewed in the CO condition (-1.647) than in the VO condition (0.607). Nevertheless, none of these values reached -/+2; that is, both were still approximately -/+1. The detrended Q-Q plots of residuals for RAU in CO vs. VO (see Appendix W for visual illustrations), respectively, exhibited two opposite patterns of skewedness – left and right. These patterns would be more extreme if we chose to examine the residuals of intelligible words rather than RAUs, but because a repeated-measures ANOVA test was performed using RAUs, there was no need to examine the

word-per-sentence distribution.

Thus, the situation in Experiment 3 mirrored the situation in Experiment 2 and the same decisions were made. I decided to run two tests: a repeated-measures ANOVA and a repeated-measures regression. The ANOVA test would reveal any statistical differences between the responses to the two types of stimuli, while the regression test would reveal whether speech in CO vs. VO can predict speech intelligibility. ANOVA uses RAUs as a continuous dependent variable with both positive and negative values, whereas regression uses the number of intelligible words as a count dependent variable with no negative values. The details and results for both tests are outlined below.

I performed a General Linear Model Repeated-Measures ANOVA test to investigate the main effect of the independent factor, consonantal vs. vocalic information, using RAUs. The model indicated a significant main effect for the factor under investigation at the 0.01 level, $F(1, 19) = 15.70$, $p= 0.001$, $\omega^2 = 0.45$. The effect size was large, as shown by the partial eta squared in the last column in Table 21, $\omega^2 = 0.45 > 0.14$. Note that because there was only one factor here, the partial eta squared provided by the software and reported here in fact equals the eta squared value, as we only have one dependent variable. Sphericity was met because the dependent factor only had two levels and hence computing χ^2 would be impossible. However, corrections are nevertheless provided in Table 21, which show no actual impact on the adjusted results.

Table 21. Summary of repeated-measures ANOVA in Experiment 3

Source		Type III Sum of Squares	*df*	Mean Square	F	Sig.	Partial Eta Squared
CO-VO	Sphericity Assumed	42383.665	1	42383.665	15.702	**.001**	.452
	Greenhouse-Geisser	42383.665	1.000	42383.665	15.702	**.001**	.452
	Huynh-Feldt	42383.665	1.000	42383.665	15.702	**.001**	.452
	Lower-bound	42383.665	1.000	42383.665	15.702	**.001**	.452
Error (COVO)	Sphericity Assumed	51287.015	19	2699.317			
	Greenhouse-Geisser	51287.015	19.000	2699.317			
	Huynh-Feldt	51287.015	19.000	2699.317			
	Lower-bound	51287.015	19.000	2699.317			

GEE was then used to perform a repeated-measures negative binomial regression analysis (see Appendix X for the model details), following the same procedure in Experiment 2. The two tables, Tables 22 and 23, provide the tests of model effects and the individual parameter estimates in Experiment 3, respectively.

Table 22. Summary of the tests of model effects in Experiment 3

Source	Wald χ^2	*df*	Sig.
(Intercept)	238.880	1	**0.001**
CO vs. VO	95.277	1	**0.001**

Dependent variable: Intelligibility (number of words)

Model: (Intercept), CO vs. VO

The model showed that CO vs. VO statistically significantly predicted intelligibility, Wald $\chi^2(1) = 95.277$, $p= 0.001$ (see Table 22 above). Since there are no other independent factors in the model, the results here are sufficient to show that CO vs. VO is a significant explanatory factor, but Table 23 provides the individual parameters and informs more about the independent factor. The table shows that speech in the CO category was significantly predicted to have a higher *log* count at 3.260 times (≈ 72.52%) more than that in the VO category, $b= 1.182$, $SE= 0.1211$, 95% CI = [0.9451, 419], Wald $\chi^2(1) = 95.277$, $p= 0.001$, Exp(B)= 3.260 (≈ 72.52%), 95% CI = [2.572, 4.134]. Table 24 below compares the results from both experiments. It shows that speech intelligibility was higher in the CO conditions than in the VO conditions in both experiments, higher in the CO condition from Experiment 3 than in the CO

condition from Experiment 2, and higher in the VO condition from Experiment 2 than in the VO condition from Experiment 3.

Table 23. Summary of the GEE parameters in Experiment 3

Parameter	B	Std. Error	95% Wald CI		Hypothesis Test			Exp(B)	95% Wald CI for Exp(B)	
			Lower	Upper	Wald χ^2	*df*	Sig.		Lower	Upper
(Intercept)	.496	.1294	.242	.749	14.686	1	**.000**	1.642	1.274	2.115
[CO vs. VO=1.00]	1.182	.1211	.945	1.419	95.277	1	**.000**	3.260	2.572	4.134
[CO vs. VO=0.00]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)										
(Negative Binomial)	4.199E-8b									

Dependent variable: Entire-Sentence Intelligibility

Model: (Intercept), CO vs. VO

Table 24. A comparison between the results in Experiments 2 and 3

Experiment	Intelligibility	CO	VO
Experiment 2	%	75.06	53.11
	RAU	77.60	52.58
	# of Words	4.5	3.19
Experiment 3	%	90.41	28.00
	RAU	91.47	27.77
	# of Words	5.34	1.65

3.3.3 Discussion and Conclusion

The main goal of Experiment 3 was to verify the findings in Experiment 2 and further investigate the role of S-Ratio. The overall results from Experiment 3 show that speech intelligibility was considerably higher in the CO condition than in the VO condition. The imbalanced S-Ratio (i.e., reduction of vowels) in this experiment improved intelligibility in the CO by approximately 15% compared to intelligibility when S-Ratio was balanced in Experiment 2. Contrariwise, speech intelligibility deteriorated in the VO condition by approximately 20%. Both S-Ratios and segmental contributions exhibit a pattern completely opposite to the pattern found in previous studies for concatenative languages such as English, as discussed below.

3.3.3.1 Segmental contributions. All results from all experiments, including Experiment 3, support the main claim that consonantal information is the primary component for speech intelligibility in Arabic. However, the disparity between participants' performance in the CO condition and the VO condition is unprecedented in the literature. The difference (62.41% and 63.70 RAU) in intelligibility rates between the CO condition and VO condition is tremendous (90.41% - 28.00% = 62.41%, 91.47 RAU – 27.77 = 63.70 RAU). Without needing to perform a one-by-one comparison with results from previous studies, we can confidently assert that the absence of vocalic information had only a marginal impact on speech intelligibility in the CO condition. If we had an unmodified condition, we would probably obtain a speech intelligibility rate value that is close to the one obtained in the CO condition here. Compare,

for example, the intelligibility rate in the CO condition (90.41%) to that in the unmodified original utterance condition (86.40%) from Cole et al. (1996). Arabic speakers' performance in the CO condition here is even slightly higher than English speakers' performance in Cole et al.'s original utterance condition. Likewise, Arabic speakers' performance in the CO condition here (91.47 RAU) is also higher than the young normal-hearing English speakers' performance in the unmodified original utterance condition (approximately 75.00 RAU) in Fogerty et al. (2012). In contrast to speech intelligibility in the CO condition, intelligibility in the VO condition was by far lower (28.00% and 27.77 in RAU) to such an extent that it seems to be the lowest value reported for a VO condition in the literature so far. This suggests that, in addition to the presence of consonantal information or lack thereof, S-Ratio played a role to increase speech intelligibility in the CO condition and decrease it in the VO condition for Experiment 3, unlike in Experiment 2. This is explored further in the next subsection.

3.3.3.2 Segmental ratios. S-Ratio in Experiment 3 was imbalanced by reducing the number of vowels by 10 segments, which triggered a greater difference not only between the values in the CO condition and the VO condition but also between the results in Experiment 2 and Experiment 3 (refer to Table 24).

It was already expected that speech intelligibility will suffer in the VO condition, simply because less vocalic information remained in the speech for the participants, which in turn reduced the magnitude of structural information that participants could use to restore some phonemes and hence identify some words. However,

the degree of speech intelligibility in the CO condition was somewhat surprising, especially given that the amount of consonantal (and hence lexical) information in the CO condition for this experiment was exactly the same as that in the CO condition from Experiment 2. There are two possible reasons for this observation. First, it could be the case that, due to comparably few occurrences of silence replacements in the CO condition in Experiment 3 as compared to Experiment 2, the speech stream was less interrupted and hence more intelligible. Second, it is possible that with fewer silence replacements participants were better able to recognize lexical boundaries and assign consonants to their original roots. It is important to note that the improved performance in Experiment 3 would not be a matter of practice effect, as the interval between the two experiments was over 6 months.

The sentences with an imbalanced S-Ratio here would be a close match to the English sentences used in previous studies, as one would predict that English sentences tend to have more vowels than consonants due to the structure of English. Nevertheless, the output in this experiment is extremely different; in fact, it is antithetical to previous findings. There is no known factor other than the distinction between English and Arabic in terms of their morphological systems to explain this. For an Arabic speaker, consonantal information is the primary conveyor of lexical information, the absence of which causes speech intelligibility to suffer. This was also observed in Experiment 1 in which S-Ratio was manipulated at the word level. The 3C-3V vs. 3C-2V S-Ratios at the word level in Experiment 1 resemble the 23C-23V vs. 23C-13V S-Ratios at the sentence level in Experiment 3, respectively. As shown in

Table 25 below, intelligibility is higher in the CO condition when S-Ratio is imbalanced compared to when it is balanced both at the word level (Experiment 1) and sentence level (Experiments 2 and 3). Unfortunately, it was impossible to include another experiment in which vowels exceed consonants at the sentence level. If it were possible to do so we would expect a result that is similar to that from the 2C-3V condition at the word level (47.50% and 47.73 RAU in the CO condition vs. 15.00% and 15.04 RAU in the VO condition) – intelligibility would still be higher in the CO condition than in the VO condition but not as high as when the S-Ratio is balanced or imbalanced to favor consonants. That is, fewer consonants and more silence replacements in the CO condition would reduce intelligibility compared to when the number of consonants is equal to or more than the number of vowels, specifically because less lexical information would be introduced in the speech. On the other hand, more vowels and fewer silence replacements in the VO condition would slightly enhance speech intelligibility compared to when the number of vowels is equal to or more than the number of consonants, since richer structural information would be introduced in the speech.

Table 25. A Comparison between the Results in Experiment 1 and Experiments 2 and 3

Experiment	Intelligibility	Balanced Ratio		Imbalanced Ratio	
		CO	VO	CO	VO
Experiment 1	%	65.00%	2.5	82.50	5.00
	RAU	63.81	- 5.36	81.94	0.08
Experiment 2 and 3	%	75.06	53.11	90.41	28.00
	RAU	77.60	52.58	91.47	27.77

In summary, the results from Experiment 3 are in concurrence with those from Experiments 1 and 2 and exhibit a great challenge to those from previous studies. In all three experiments speech intelligibility was consistently greater when Arabic participants were presented with speech containing only consonants and consistently less when participants were presented with speech containing only vowels. Such findings agree with previous findings at the word level, but completely contradict previous results at the sentence level, a conclusion that can be attributed to the difference between the morphological systems of concatenative languages and nonconcatenative languages. This was further supported by the disparities between the intelligibility rates associated with different S-Ratios. To emphasize this morphology-based distinction between the two categories of languages and to examine the role of S-Ratio in this context, the next chapter will re-visit segmental contributions in English, comparing balanced and imbalanced ratios in the language with two further experiments – Experiments 4 and 5.

⇐
Chapter
Three

CHAPTER FOUR:

⇒
Chapter
Five

SEGMENTAL CONTRIBUTIONS AND RATIOS IN ENGLISH

The first three experiments from Chapter 3 examined the contributions and effects of segments and S-Ratios on speech intelligibility in Arabic. The overall results showed that consonantal information contributed more to speech intelligibility than vocalic information at both the word level and sentence level, and that S-Ratio was a critical factor at both levels.

This chapter will investigate the role of segments and S-Ratio in concatenative languages, specifically in English. I hypothesize that the role of S-Ratio in concatenative languages is different from its role in nonconcatenative languages. Specifically, because consonants are privileged in the morphological system of Arabic but not in English, we should expect a greater impact on speech intelligibility from a change in S-Ratios in Arabic as compared to in English. To examine this hypothesis, I conducted two experiments in which I manipulated S-Ratios within both words and sentences in a manner similar to the previous experiments, but with minor changes as necessitated by the English morphology.

Experiment 4 is designed to test balanced vs. imbalanced S-Ratios in individual English words in order to test whether both vocalic and consonantal information play similar roles in speech intelligibility. Experiment 5 is similar to Experiment 4 but examines S-Ratios in English sentences rather than words. Both experiments were meant to show how manipulating the ratio of segments will exert a different

effect on the CO condition compared to the VO condition. Hence, this will show further evidence for the overall claim being investigated in this dissertation regarding the difference between nonconcatenative and concatenative languages.

4.1 Experiment 4

The methodological approach in Experiment 4 followed that in Experiment 1, with some modifications relevant to the English language.

4.1.1 Methodology

4.1.1.1 Overview. I selected 48 English words, half of which have a balanced ratio of consonants and vowels and half of which have an imbalanced ratio. A native English speaker recorded the words, and the same silence replacement method was used to generate two versions of stimuli, VO and CO, for each ratio type.

4.1.1.2 Word length. Of the 48 words (see Appendix Y for full list), half had a balanced number of vowels and consonants, 2C-2V, 3C-3V, and 4C-4V as in "goofy" /gufi/, "jealousy" /ʤɛləsi/, and "delicacy" /dɛləkəsi/, respectively, and the other half had an imbalanced number, 2C-1V, 3C-1V, and 4C-1 as in "vet" /vɛt/, "gift" /gɪft/, and "trench" /trɛnʧ/, respectively. The average length of words was 5.16 segments, compared to 5.33 in the Arabic stimuli of Experiment 1.

4.1.1.3 Word selection. As was the case for the Arabic target words, I chose words that are familiar to English speakers and that have a variety of different vowels and consonants. Word frequency will be outlined in Subsection 4.1.1.4, but there were several other considerations in selecting the words. First, I thought of the initial

segment (vowel vs. consonant) as an important factor in word recognition and controlled for it, at least in the balanced ratio stimuli. Every word with an initial consonant and a balanced ratio had a counterpart word with an initial vowel, such as "goofy" /gufi/ vs. "edit" /ɛdət/, "modify" /mɑdəfɑɪ/ vs. "elevate" /ɛləveɪt/, and "delicacy" /dɛləkəsi/ vs. "analytic" /ænəlɪtɪk/. However, in the imbalanced ratio condition, it was impossible to contrast initial segments in this manner, because although a there are some words with VCC structure such as "act" and a few with VCCC such as "angst", words with the structure VCCCC do not exist in English. This contrast between consonants and vowels in word initial position was not initially meant as a major predictor variable in the experiment design, but it was included in order to avoid any bias in the VO vs. CO conditions, as the initial segment may play a decisive role in word recognition. The intention was to check if this factor would affect speech intelligibility. If not, words starting with vowels would be grouped with words starting with consonants.

Second, within vowels and consonants, I also tried to have as many different segments as possible (e.g., stops vs. fricatives vs. approximants, voiced vs. voiceless consonants, high vs. low vowels, and front vs. back vowels). Third, in the imbalanced condition, I took into consideration the distribution of segments to avoid having clusters in only one position. For example, 4C-1V words had a CVCCC structure as in "tempt" /tɛmpt/, a CCVCC structure as in "trench" /trɛnʧ/, and a CCCVC structure as in "strait" /streɪt/. Fourth, I also tried to have a diversity of consonants within each

cluster and avoided having one type of combination. Initial clusters consisted of segments such as /θr/, /sk/, /sp/, /tr/, /ʃr/, /sl/, /bl, /dr/, /str/, and /skr/ while coda clusters consisted of segments such as /ft/, /kt/, /nk/, /lb/, and /mpt/ (see Appendix Y).

4.1.1.4 Frequency and judgements. There were already several constraints on selecting isolated words, but I also controlled for word frequency to reach a similar level of frequency in both the balanced ratio and imbalanced ratio conditions and within each condition. To match frequencies, I used The Corpus of Contemporary American English (COCA) (Davies, 2017). Table 26 below summarizes the means and standard deviations of word frequency for all conditions.

Table 26. Word frequencies in Experiment 4

	Balanced with Initial C			Balanced with Initial V			Imbalanced		
	4C-4V	3C-3V	2C-2V	4C-4V	3C-3V	2C-2V	4C-1V	3C-1V	2C-1V
M	3130	3111	3016	3056	3040	3058	3020	3092	3038
STD	(1233)	(1262)	(1258)	(1237)	(1264)	(1270)	(1205)	(1286)	(1285)
M	3086			3051			3050		
STD	(1133)			(1137)			(1139)		
M	3068						3050		
STD	(1118)						(1139)		

Note: the total number of words in the corpus is 560 million in 2017. Hence, frequency here is per million.

Although balancing the frequencies should eliminate any concerns about the participants' familiarity with the lexical items used in this experiment, I also had the words judged by native speakers of English. I followed the familiarity judgement and criteria used for the Arabic stimuli to make certain that English speakers are familiar with the word list just as Arabic speakers were in Experiment 1. The sentences were sent to 10 judges in written from and later to 5 of them in spoken form for familiarity judgements. The final version of words passed the criterion with a 6.70 in written form and 6.21 in spoken form (see Appendix Z for judgement rates).

4.1.1.5 Recording and stimulus preparation. After I finalized the list of words, I had a male native speaker of English record the stimuli. He was given the same instructions that the Arabic speaker was given in Experiments 1, 2, and 3. I then used the same techniques as with the previous stimuli to generate two conditions for each word: VO and CO. This yielded four levels of treatments as shown in Table 27 below.

Table 27. Illustration of the English word experiment design

Balanced Ratio		Imbalanced Ratio	
VO	CO	VO	CO
/_u_i/ "goofy"	/g_f_/ "goofy"	/_ɛ_/ "vet"	/v_t/ "vet"
2C-2V		**2C-1V**	

4.1.1.6 Task. Previous studies using English stimuli asked the participants to write down what they could hear, a technique that was inappropriate in the case of

Arabic for reasons outlined in Chapter 3. For the current experiment, I asked the participants to record what they heard using verbal reporting, to be consistent the previous three experiments on Arabic.

4.1.1.7 Participants. Twenty native English speakers participated in the experiment for a compensation of $10 each. None of them had exposure to the stimuli during any phase of word selection, rating, or recording. Participants were undergraduate or graduate students (or their relatives) at UW-Milwaukee. Almost half of them (9) were male and almost half were female (11), with a 18-40 range of age (M= 26.15, STD = 7.45). One additional subject participated in the experiment but was eventually removed from the sample, due to the observation that she did not identify any word correctly. This subject also participated in Experiment 5 and identified only a few words from 48 sentences and was removed from that data set, as well.

4.1.1.8 Procedure. The same procedures outlined in the previous experiments were followed in Experiment 4. Each participant listened to 48 words and was asked to verbally report what they heard while being recorded. The order of words alternated between CO and VO conditions, between balanced and imbalanced conditions, and between initial C and Initial V conditions. No two words from the same condition were adjacent in the stimuli presentation nor did any participant listen to the same word in two different conditions.

4.1.1.9 Data scoring and analysis. The responses were scored based on their accuracy and assigned a percentage score as in Experiment 1. The overall scores for each condition were then transferred to RAU for comparison with previous findings

in the literature. This should also make the comparison between the results from this experiment and those from Experiment 1 more straightforward.

4.1.2 Results.

The results from this experiment show that, in general, speech intelligibility was higher in the CO condition than in the VO condition regardless of S-Ratio, except when the initial segment was a vowel (see Figure 10 and Table 28 below for the results in percentage and Appendix A2 for results in RAU). This may appear to mean that consonants play a greater role in speech intelligibility than vowels do at the word level, which would be an inaccurate conclusion. In fact, speech intelligibility was almost equal in the CO and VO conditions, but the effect of initial segment led to this outcome. Therefore, we should approach the results from a different angle and examine the results away from the initial segment effect.

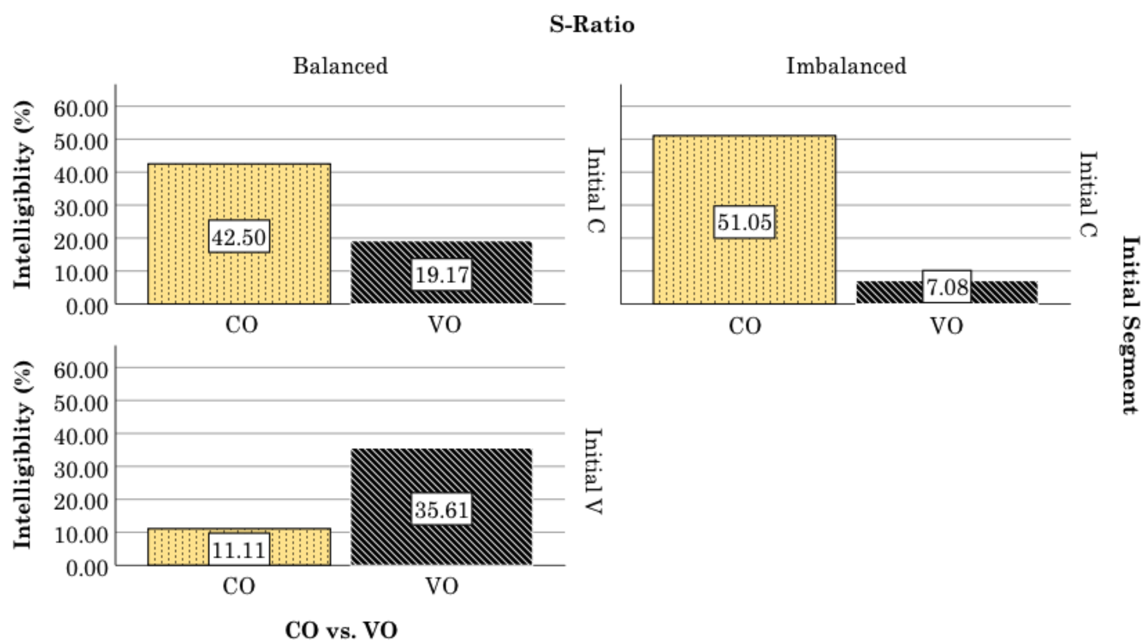


Figure 10. Intelligibility (%) broken down by C0 vs. VO, Initial C vs. V & S-Ratio in Experiment 4

The two types of initial segment showed an opposite pattern when placed in the CO condition vs. VO condition, and they should cancel each other if we collapse them into one group under the balanced ratio. As shown in Table 28 below, grouping the results from the two conditions (Initial C vs. Initial V) for the balanced ratio *did* provide us with roughly the same intelligibility rate in the CO condition (53.61%) vs. VO condition (54.78%). The overall intelligibility rates in the balanced ratio (54.19% [*M*= 27.09%]) and in the imbalanced ratio (58.13% [*M*= 29.06%]) became very close (see Appendix B2 for visual presentations). When we increased consonantal information and reduced vocalic information in the imbalanced ratio, we obtained a higher intelligibility rate for consonants, a lower intelligibility rate for vowels, and a large difference between the two conditions (51.05% – 7.08%= 43.97%). When we balanced the consonantal information and vocalic information in the balanced ratio, we obtained similar intelligibility rates and an appreciably small difference (54.78%-53.61%= 1.17%) between the two conditions.

Table 28. Results grouped into two conditions (Balanced vs. Imbalanced) in Experiment 4

CO vs. VO	Balanced Ratio		Imbalanced Ratio
	Initial C	Initial V	
CO	42.50	11.11	51.05%
VO	19.17%	35.61%	7.08%
Total CO	53.61%		51.05%
Total VO	54.78%		7.08%
Total per S-Ratio	54.19		58.13
Means	25.80	27.39	29.06
	27.09		

The current experiment also revealed novel findings that did not emerge in previous studies on English. In general, the results show a consonantal privilege when the initial segment is a consonant, and this was true regardless of whether the S-Ratio is balanced (42.50%, 43.06 RAU in the CO condition and 19.17%, 19.42% RAU in the VO condition) or imbalanced (51.05%, 51.43 RAU in the CO condition and 7.08%, 2.35 RAU in the VO condition), and show a vocalic privilege when the initial segment is a vowel (11.11%, 9.05 RAU in the CO condition, and 35.51%, 36.55 RAU in the VO condition).

For statistical analysis, I followed the procedure in Experiment 1. I used a GEE repeated-measures logistic regression model (see Appendix C2 for model information) to test for the main effects of and interactions between predictors. The predictor variables were CO vs. VO, Initial Segment (Initial C vs. Initial V), and S-Ratios (note that the six types of ratios were grouped into two categories: Balanced vs. Imbalanced). The intelligibility score (0 or 1) was set up as a binary dependent variable. Subject and trial were specified as a subject variable and repeated effect variable, respectively. The goodness of fit (Appendix C2) showed that *independent*, as opposed to *exchangeable* in Experiment 1, as the working correlation matrix was the best (QIC = 997.561 and QICC = 994.158), although the difference between *exchangeable* and *independent* was de facto minor (QIC = 998.025 and QICC = 994.349 for *exchangeable*). This simply means that the responses from a single subject were not truly dependent/correlated, which would allow us to build a normal (non-repeated-measures) binominal logistic regression model and obtain a similar output. Notwithstanding,

with *independent* as a working correlation matrix, the software should adjust itself when it finds some instances of correlation while performing different iterations, a reason for favoring this option over others. In addition, I took a step further and ran the same analysis with *exchangeable* as the working correlation matrix once and a normal (non-repeated-measures) binominal logistic regression once, and the results were almost identical to those when using *independent*.

The GEE model had three main effect factors (CO vs. VO, S-Ratio, & Initial Segment) but only two two-way interaction effects (CO vs. VO x Initial Segment, and CO vs. VO x S-Ratio). The Initial Segment x S-Ratio interaction was deliberately removed from the model, because one level of the S-Ratio (i.e., Imbalanced) had only one level of Initial Segment (i.e., Initial C) due to some experimental design constraints. It was possible to keep all two-way interactions to include this one, because the model would reach one iteration where it decides that computing the interaction between the two factors would be impossible and then dismiss it from the model. However, I removed this interaction manually to avoid the model having to perform any unnecessary equations.

The model was able to predict overall 72.20% (693 cases) of the responses correctly but misclassified 27.8% (267 cases) of the responses (Appendix C2). As shown in Table 29, the model showed that CO vs. VO significantly predicted whether speech will be in the intelligible vs. unintelligible category, Wald $\chi^2(1) = 5.405$, $p= 0.02$., and so did the S-Ratio, Wald $\chi^2(1) = 5.229$, $p= 0.022$. However, the Initial Segment did not statistically significantly predict whether speech will be intelligible or unintelligible,

Wald $\chi^2(1) = 3.186$, $p= 0.74$. However, the Initial Segment significantly interacted with CO vs. VO; intelligibility of speech with Initial Segment depended on whether it was in the CO or VO condition, Wald $\chi^2(1) = 48.562$, $p= 0.001$. The CO vs. VO factor also interacted with the S-Ratio factor, which means that the predictability of intelligibility category for speech with a specific S-Ratio was dependent on whether speech was in the CO condition or in the VO condition, Wald $\chi^2(2) = 12.395$, $p= 0.001$.

Table 29. Summary of tests of the model effects in Experiment 4

Source	Wald χ^2	*df*	Sig.
(Intercept)	73.933	1	**.000**
CO vs. VO	5.405	1	**.020**
Initial Segment	3.186	1	.074
S-Ratio	5.229	1	**.022**
CO vs. VO x Initial Segment	48.562	1	**.000**
CO vs. VO x S-Ratio	12.395	1	**.000**

Dependent variable: Intelligibility

Model: (Intercept), CO vs. VO, Initial C vs. Initial V, S-Ratio, CO vs. VO x Initial C vs. Initial V, CO vs. VO x S-Ratio

The model also produced the individual parameters as provided in Table 30, which shows both the individual parameter estimates and the odds ratios. As was decided in Experiment 1, another Table (Table 31) was produced with *unintelligible* as the reference category to make it easier to understand the likelihood when

Exp(B)<1 and/or $b < 0$. The two tables provide only the significant and non-redundant results (see Appendix D2 for complete tables). The first individual parameter shows that speech in the CO category was significantly predicted to be in the intelligible category with 4.424 times (≈ 81.56%) higher likelihood than when speech was in the VO category, b= 1.487, SE= 0.3034, 95% CI = [0.892, 2.082], Wald $\chi^2(1) = 24.020$, p= 0.001, Exp(B)= 4.424 (≈ 81.56%), 95% CI = [2.441, 8.017]. Although Initial Segment, which was only available in a portion of the data as will be discussed later, was not a statistical predictor in the overall model, Initial C turned out to be a predictor factor. Speech from the Initial C category was 2.332 (≈ 69.99%) more likely to be in the intelligible category than that from the V category, b= 0.847, SE= 0.2053, 95% CI = [0.444, 1.249], Wald $\chi^2(1) = 17.002$, p= 0.001, Exp(B)= 2.332 (≈ 69.99%), 95% CI = [1.559, 3.487]. Speech that had a balanced ratio was significantly 3.11 times (≈ 75.67%) more likely to fall in the intelligible category than speech with an imbalanced ratio, b= 1.135, SE= 0.2147,95% CI = [0.714, 1.556], Wald $\chi^2(1) = 27.992$, p= 0.001, Exp(B)= 3.11 (≈ 75.67%), 95% CI = [2.042, 4.738].

The model significantly predicted that the likelihood of speech from the Initial C category being in the intelligible category depended on being in the CO category 13.789 times (≈ 93.24%) more than the counterpart category, b= 2.624, SE = 0.3765, 95% CI = [1.886, 3.362], Wald $\chi^2(1) = 48.562$, p= 0.001, Exp(B)= 13.789 (≈ 93.24%), 95% CI = [6.592, 28.843]. The model also significantly predicted that the likelihood of speech in the balanced ratio category being in the unintelligible (rather than intelligible) category was dependent on being in the CO category 4.388 times (≈ 81.44%)

more than in other categories, b= 1.479, SE = 0.4201, 95% CI = [0.656, 2.302], Wald $\chi^2(1)$ = 12.395, p= 0.001, Exp(B)= 4.388 (≈ 81.44%), 95% CI = [1.926, 9.996]

Table 30. Summary of the GEE parameters in Experiment 4 (Ref. Category= Intelligible)

Parameter [a]	B	Std. Error	95% Wald CI		Hypothesis Test			Exp(B)	95% Wald CI for Exp(B)	
			Lower	Upper	Wald χ^2	*df*	Sig.		Lower	Upper
(Intercept)	.593	.1578	.283	.902	14.103	1	**.000**	1.809	1.327	2.464
[CO vs. VO=1][b]	1.487	.3034	.892	2.082	24.020	1	**.000**	4.424	2.441	8.017
[Initial Segment=1]	.847	.2053	.444	1.249	17.002	1	**.000**	2.332	1.559	3.487
[S-Ratio=1]	1.135	.2147	.714	1.556	27.922	1	**.000**	3.110	2.042	4.738
[CO vs. VO=1] x [Initial Segment=1]	2.624	.3765	1.886	3.362	48.562	1	**.000**	13.789	6.592	28.843
[CO vs. VO=1] x [S-Ratio=1]	-1.479	.4201	-2.302	-.656	12.395	1	**.000**	.228	.100	.519

Dependent variable: Intelligibility

Model: (Intercept), CO vs. VO, Initial Segment, S-Ratio, CO vs. VO x Initial Segment, CO vs. VO x S-Ratio, N vs. V x S-Ratio

a. Reference category: *Intelligible.*

b. CO is coded as 1, VO as 0, Initial C as 1, Initial V as 0, Balanced as 1, and Imbalanced as 0

Table 31. Summary of the GEE parameters in Experiment 4 (Ref. Category= Unintelligible)

Parameter [a]	B	Std. Error	95% Wald CI		Hypothesis Test			Exp(B)	95% Wald CI for Exp(B)	
			Lower	Upper	Wald χ^2	*df*	Sig.		Lower	Upper
(Intercept)	-.593	.1578	-.902	-.283	14.103	1	**.000**	.553	.406	.753
[CO vs. VO=1][b]	-1.487	.3034	-2.082	-.892	24.020	1	**.000**	.226	.125	.410
[Initial Segment=1]	-.847	.2053	-1.249	-.444	17.002	1	**.000**	.429	.287	.641
[S-Ratio=1]	-1.135	.2147	-1.556	-.714	27.922	1	**.000**	.322	.211	.490
[CO vs. VO=1] x [Initial Segment=1]	-2.624	.3765	-3.362	-1.886	48.562	1	**.000**	.073	.035	.152
[CO vs. VO=1] x [S-Ratio=1]	1.479	.4201	.656	2.302	12.395	1	**.000**	4.388	1.926	9.996

Dependent variable: Intelligibility

Model: (Intercept), CO vs. VO, Initial Segment, S-Ratio, CO vs. VO x Initial Segment, CO vs. VO x S-Ratio, N vs. V x S-Ratio

a. Reference category: *Unintelligible*

b. CO is coded as 1, VO as 0, Initial C as 1, Initial V as 0, Balanced as 1, and Imbalanced as 0

Two potential concerns are associated with our data and results. First, we should recall that the S-Ratio has only two levels, Balanced and Imbalanced, but each has three sub-levels of ratios, depending on the number of consonants vs. vowels. I purposely opted not to fit these sub-levels in the model for three reasons. First, there was no research question or hypothesis about the differences between the ratio types within the balanced or imbalanced condition. Second, these six levels would add tens of interactions with other factors and make it overwhelmingly difficult to interpret the results and draw a conclusion about the effect of balanced vs. imbalanced ratio on speech intelligibility. Third, the pattern of intelligibility for each sub-level was consistent with other sub-levels of the same category, as illustrated in Figure 11. However, as already alluded to, there is a concern about the absence of Initial V in the Imbalanced condition, not only because we could not test the role of initial segment in that condition, but also because it led to a disparity in the number of trials per condition (a sampling size issue). Specifically, although there are an even number of trials in the Balanced vs. Imbalanced conditions, only half of the stimuli in the Balanced condition correspond to those in the Imbalanced condition (24 in Balanced vs. 48 in Imbalanced). This is potentially why the Initial Segment factor did not reach statistical significance, since only one fourth of the data (25.00% = 24 out of 96 trials per subject) was in the Initial V condition.

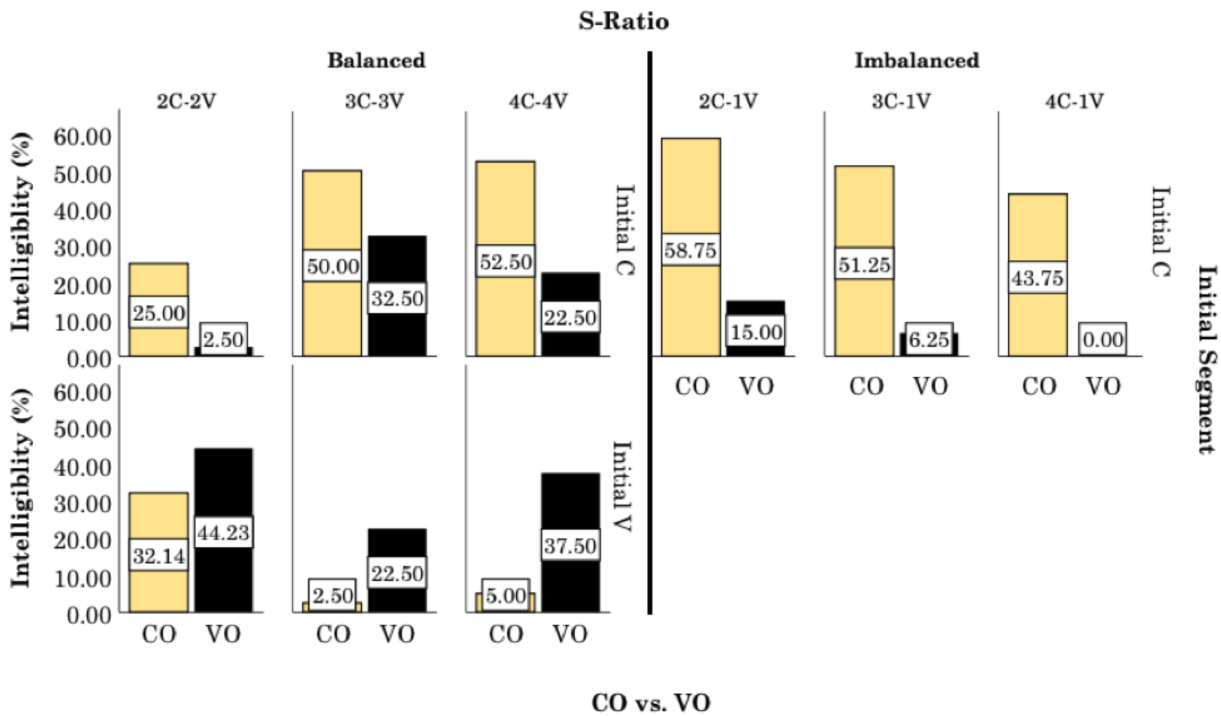


Figure 11. Intelligibility (%) broken down by CO vs. VO, Initial C vs. V & S-Ratios in Experiment 4

The concern articulated above becomes more apparent when looking at the reverse patterns of intelligibility in Initial C vs. Initial V for the Balanced condition. Figure 12 demonstrates this difference by grouping all sub-levels into one category. It is clear that speech intelligibility proportions show an asymmetric pattern that deserves further investigation. There are 12 types of trials in the Balanced condition (3 [2C-2V vs. 3C-3V vs. 4C-4V] x 2 [CO vs. VO] x 2 [Initial C vs. Initial V]) and 6 types in the Imbalanced condition (3 [2C-1V vs. 3C-1V vs. 4C-1V] x 2 [CO vs. VO]), making up a total of 18 types. I ran a macro[16] to calculate the intelligibility average of each type for each subject, following the recommendation made by Max and Onghena

[16] The macro script is provided in Appendix E2.

(1999). I then transformed all averages to RAUs, using two different default numbers of trials for the Balanced (4 per sub-type) vs. Imbalanced (8 per sub-type) conditions in preparation for an ANOVA test. The ANOVA test was performed in order to eliminate the potential concerns with our data, as discussed above.

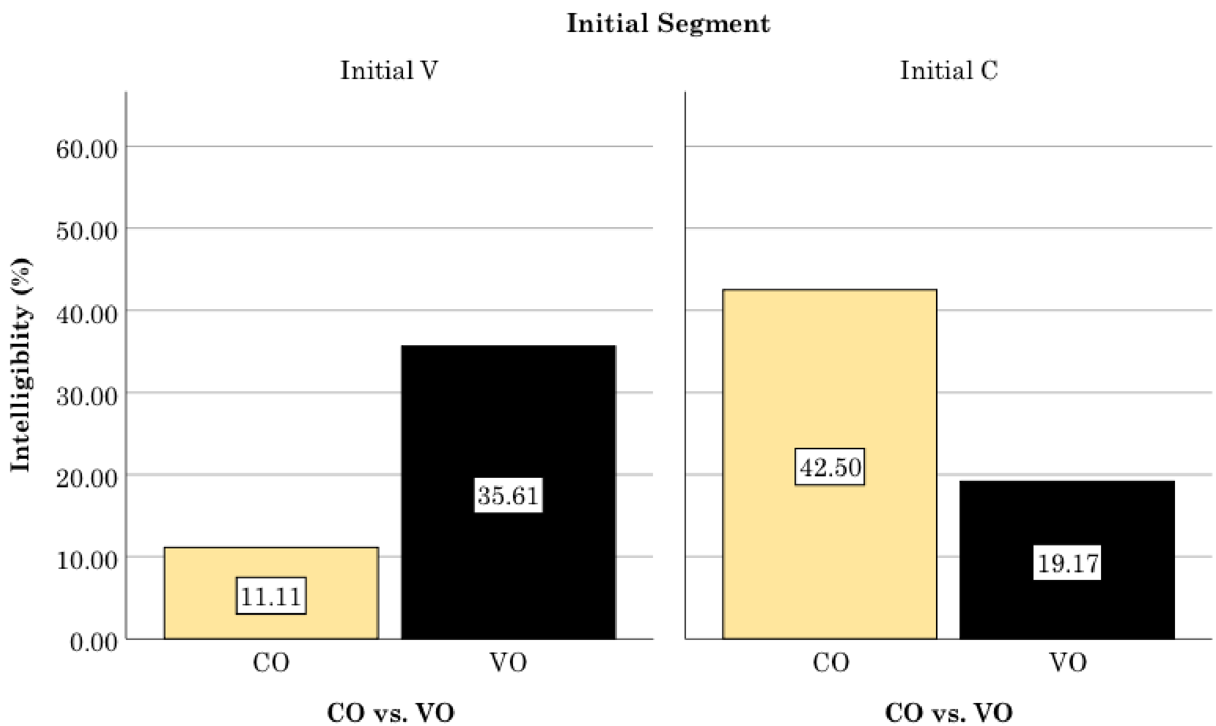


Figure 12. Intelligibility (%) broken down by CO vs. VO, Initial C vs. V in Experiment 4

There were two options for an ANOVA test: univariate (because we only have one dependent factor) or repeated-measures (because multiple responses came from one participant). The previous GEE test showed that responses from each individual participant were not clearly dependent/correlated, which would allow us to run a non-repeated-measures univariate ANOVA. However, I dismissed this option, and pro-

ceeded with the second one. It was important to check for the repeated-measures assumptions (sphericity and normality) first. The sphericity test was important only for the S-Ratio having three levels, but for the other two factors having only two levels each, Initial C vs. Initial V and CO vs. VO, it was not needed. Mauchly's test of sphericity indicated that the assumption of sphericity was not violated for S-Ratio in the Balanced condition, $\chi^2(2) = 3.382$, $p= 0.184$, or in the Imbalanced Condition, $\chi^2(2) = 2.696$, $p= 0.260$. Normality was, however, violated, as was shown by a Shapiro-Wilk normality test, $p=0.05$, but both kurtosis and skewedness tests showed low values that centered around -/+1, supporting the decision to run the test therewith. The conservative Bonferroni correction was also used to adjust the confidence interval for the multiple comparison tests.

The ANOVA test indicated no statistically significant differences between the S-Ratios within the Balanced condition (2C-2V vs. 3C-3V vs. 4C-4V) or within the Imbalanced condition (2C-1V vs. 3C-1V vs. 4C-1V) at 0.05 level. This supports the previous decision to group these S-Ratios under two categories: Balanced and Imbalanced. However, and more importantly, the test *did* indicate a significant main effect for the Initial Segment factor at the 0.01 level (see Appendix F2 for all comparisons). There was a statistically significant difference in intelligibility for Initial C vs. Initial V in the Balanced condition, $F(1, 19) = 9593$, $p= 0.006$, with a large effect size $\omega^2 = 0.33>0.14$. This result upholds the questionability of the results for the Initial Segment factor in the GEE test, and justifies the decision for running another test to check for the differences produced by the Initial Segment factor.

Second, absence of Initial Segment factor in the Imbalanced condition did not allow us to explore interaction between the two factors. It also affected our understanding of the overall difference between the Balanced and Imbalanced conditions, because Initial V yielded a pattern that made the comparison hazy. Therefore, the solution was to either collapse the data from Initial C and Initial V into one category, which would have resulted in an equal size of samples, or to remove Initial V from the data and then run a repeated-measures ANOVA. The decision was to run a repeated-measures ANOVA without including the responses to Initial V stimuli in the data (see Appendix F2 for the summary table). The sample size was equal in both conditions (CO vs. VO and S-Ratio [Balanced vs. Imbalanced), sphericity was assumed since we only had two levels, and the kurtosis and skewedness values were low (about -/+1). There was a statistically significant difference in intelligibility for CO vs. VO, $F(1, 19) = 114.895$, $p= 0.001$, with a large effect size $\omega^2 = 0.85>0.14$. There was no statistically significant difference for the S-Ratio, $F(1, 19) = 0.214$, $p= 0.64$, or for the CO vs. VO x S-Ratio interaction, $F(1, 19) = 4.149$, $p= 0.54$.

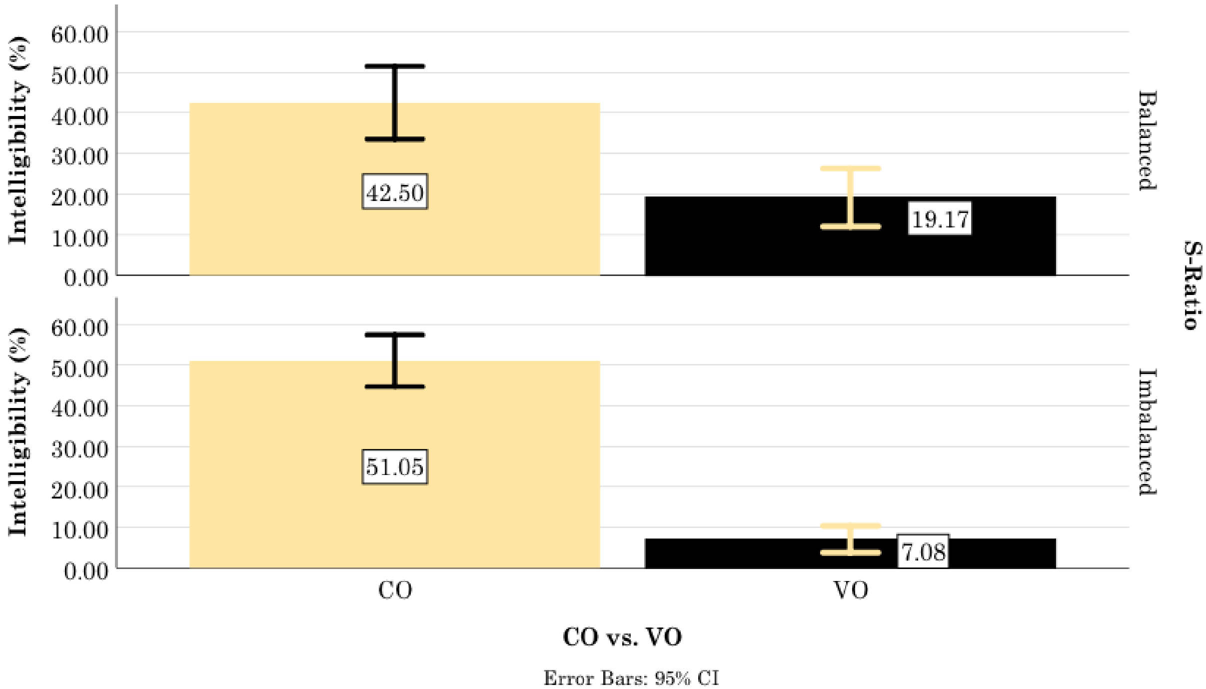


Figure 13. Intelligibility (%) broken down by CO vs. VO and S-Ratio in Experiment 4

In summary, both statistical tests showed that CO vs. VO was a major factor as a main effect and interaction effect with other factors. The GEE model showed that Initial C was an explanatory factor as a main effect and interaction effect, but Initial Segment as a general factor did not contribute to the data fit. However, the ANOVA test showed that, without the impact of the Initial Segment, there was no statistical difference between the Balanced and Imbalanced conditions. The ANOVA test did also reveal a statistically significant difference between Initial C and Initial V. This is a novel finding, as will be discussed in the subsequent section.

4.1.3 Discussion and Conclusion

Experiment 4 was designed to examine the effect of S-Ratio on segmental contribution and hence on speech intelligibility at the word level. The overall results showed that vowels and consonants played an almost equal role in speech intelligibility, and that S-Ratio had an effect on both types of segments. In addition, the experimental design was successful to obtain a new finding. Namely, vocalic information better facilitated speech intelligibility when S-Ratio was balanced, and initial segment was a vowel. This finding calls for revisiting the previous findings in the literature on concatenative languages such as English. Hence, any discussion of segmental contributions and S-Ratios cannot be dissociated from the initial segment factor here.

4.1.3.1 Segmental contributions. The participants showed better performance (i.e., higher intelligibility) when they were presented with consonantal information for lexical items starting with a consonant, regardless of whether segments were balanced or imbalanced. This finding supports what has been reported in the literature about segmental contribution in isolated words (see e.g., Cutler et al., 2000; Fogerty et al., 2012; van Ooijen, 1996; Owren and Cardillo, 2006). However, a critical question that manifests itself to previous studies is whether consonants contributed to intelligibility more than vowels did because listeners found consonants to be more informative to them or because, in earlier studies, most of the stimuli items contained a consonant in initial position. In other words, it is well known that the initial segment has a great impact on overall intelligibility (Eberhard, Spivey-Knowlton, Sedivy, &

Tanenhaus, 1995; Marslen-Wilson, 1987; Miller & Eimas, 1995), and when the initial segment is a consonant, intelligibility will be very likely to suffer when speech is presented in the VO condition. The CO condition will have a notable advantage, because the initial segment will activate candidates that include the target word. The VO condition with an initial consonant will of course lack this advantage, as the first segment the listener will encounter is not really the initial segment of the word. This may not only affect the listener's ability to recognize the target word but also mislead them to a totally different word, since the first segment they recognize may activate a completely unintended set of candidates.

This initial segment effect was observed by Owren and Cardillo (2006) who examined the effect of initial phoneme on participants performance in a portion of their stimuli (in Experiment 2) and found an identical pattern to the one reported in Figure 12 here. That is, they found superior intelligibility when the target word contained an initial vowel and the stimulus was presented in VO condition. This means that whenever the modified stimulus (CO or VO) preserves the same initial segment of the original stimulus, it will be more intelligible than when it does not. Looking at the balanced ratio panel in Figure 13 above, we can see that the difference in speech intelligibility between the CO condition and VO condition when the initial segment was a consonant (42.50 – 19.17= 23.33) is very close to that when the initial segment was a vowel (35.61 – 11.11= 24.5). Therefore, there was an advantage for Initial C once and for Initial V once, depending on the match (or mismatch) between the original initial segment and that in the modified version (CO vs. VO).

Now comes the question: does consonantal information contribute more to speech intelligibility than vocalic information does in English? Despite the fact that the figures above seem to depict a real advantage for CO over VO condition, based on our results here coupled with the brief discussion in Owren and Cardillo (2006), the answer is nevertheless likely to be *no*. When the initial segment presented to the listener in the stimulus was different from the original initial segment in the word, speech intelligibility was low, but similar in both CO and VO conditions. If one type of segment were truly superior to its counterpart, it should have led to a higher intelligibility when all other design aspects were factored out, but this was not borne out. As a matter of fact, it is possible that many previous findings in the literature were driven by this confounding factor. The 148 materials used in Fogerty et al. (2012) (as well as some other studies such as Fogerty et al. 2010) were all CVC monosyllabic words. This means that all stimuli presented in the CO condition privileged the presence of initial segment while all stimuli presented in the VO condition lacked the advantage of initial segment. This would be supported by findings reported in Fogerty et al. (2010) about the identification of initial segment vs. final segment in the CO vs. VO conditions. Fogerty et al. (2010) found that, in the CO condition, the initial C was identified better than the final C in the CO condition while in the VO condition, the final C was identified better than the initial C. Furthermore, Owren and Cardillo (2006) reported a better performance represented by a higher *d'* score for initial-vowel stimuli presented in the VO condition than for initial-consonant stimuli presented in

the CO condition (Experiment 3). This outcome was also observed in the current experiment (repeated here: 19.17% for Initial C presented in VO vs. 11.11 for Initial V presented in CO). This disparity did not reach a statistical difference in our experiment, but it did in Owren and Cardillo (2006). Their interpretation was based on the evident disparity in the number of content words starting with consonants vs. vowels. That is, during the process of lexical access and selection, the set of candidates starting with vowels were fewer than those starting with consonants, leading to better intelligibility in the VO condition for words with initial vowels.

Thus, we can posit that, for concatenative languages, a segment is a segment. Vocalic information contributes to speech intelligibility at word level as much as consonantal information does. Other factors such as word frequency, morphological structure, S-Ratio, and initial segment may influence intelligibility in one condition or another. The last factor here (initial segment) may pose the same challenge to our conclusion from Experiment 1 about the effect of the Arabic morphological system on speech intelligibility, since all words started with consonants due to some language structure constraints. However, the outcome of Experiments 2 and 3 largely ruled out this concern. This will be left for a more detailed discussion in Chapter 5.

4.1.3.2 Segmental ratios. Speech intelligibility was greater in the CO condition than in the VO condition when the initial segment was a consonant, regardless of the S-Ratio type. This may deceive us and lead to the conclusion that consonants play a greater role in speech intelligibility than vowels do at the word level. In fact, although all factors were seemingly controlled for, it is likely the outcome appeared as it did

due to the effect of initial segment as discussed earlier. Therefore, grouping the results from the Initial C and Initial V conditions was a solution to examine the effect of S-Ratio away from the initial segment effect.

The outcome after collapsing the results from both conditions indicated that vowels and consonants together contributed the same amount of information to speech intelligibility at the word level in English (as we saw in Table 28). This equal contribution of segments may be challenged by the discrepancy between the results from Fogerty et al. (2012) and the results from the 2C-1V condition here. This condition shows a higher rate of intelligibility in the CO condition, but Fogerty et al. (2012) found no statistical difference for the intelligibility rate of CVC words between the CO and VO conditions. There are two possible interpretations for this disproportion of intelligibility rates. First, the 2C-1V words were not all of CVC structure; rather, some items were of CCV structure, which probably boosted speech intelligibility in the CO condition or lowered it in the VO condition. With two adjacent consonants available to the listener in the CO condition, the pool of lexical candidates becomes smaller and the word approaches its uniqueness point more than when the listener is presented with a C followed by a silence and then another C. Second, we would expect the CVC vocalic information to bear transitional information from the two consonants, and the CCV vocalic information to bear more information from the second consonant than from the first. This could make a CVC word more intelligible than a CCV word in the VO condition, especially if we recall the initial segment effect. One evidence for this possibility comes from Fogerty et al. (2010) who scored phoneme

identification for CVC words in the CO vs. VO condition and found that "missing consonants were identified significantly better than missing vowels" (p. 3318) regardless of whether the target word was intelligible or not.

Hence, irrespective of the minor issues discussed above, our overarching results still support the hypothesis that consonants and vowels in concatenative languages are both two important acoustic sources that contribute to speech intelligibility equally. This is also predicted to be the case at the sentence level, which is the goal of Experiment 5.

4.2 Experiment 5

4.2.1 Methodology

4.2.1.1 Overview. I constructed 48 English sentences with different S-Ratios, half of which had a balanced ratio (see Appendix G2) and half of which had an imbalanced ratio of consonants and vowels (see Appendix H2). The same native English speaker from Experiment 4 recorded the sentences. The same silence replacement method was applied, and the same entire-sentence scoring method in Experiment 3 was implemented.

4.2.1.2 Sentence length. The stimuli consisted of 48 sentences. Half of the sentences had a balanced ratio (mostly 10 consonants vs. 10 vowels), while the other half had an imbalanced ratio (mostly 15 consonants vs. 5 vowels). Some sentences had an imbalanced ratio of 6:16 or a balanced ratio of 11:11, but the difference was always 10 segments (see Appendix S for details). The average number of words in all

sentences (*M*= 5.7 words overall in balanced and imbalanced conditions) was kept approximately the same as in the previous Arabic sentences (*M*= 5.8 words overall in balanced and imbalanced conditions). Given that the morphology of English is comparatively less rich, six English words forming a complete sentence had overall fewer morphemes (and hence fewer segments) in comparison to six Arabic words bearing agreement, case, number, and gender morphemes. In addition, the structure of English does not allow for the construction of long sentences with a balanced ratio because the probability of having more consonants than vowels increases as a sentence becomes longer. Nevertheless, the difference in number of segments for balanced and imbalanced ratios was the same as in the Arabic sentences (a 10-segment difference).

4.2.1.3 Sentence construction. As was the case in the previous experiments, it was necessary to construct sentences to meet some design requirements such as sentence length and S-Ratios, but also possible to adopt, with minor amendments, some sentences from online resources such as news outlets. The English sentences with a balanced ratio were comparable to the Arabic sentences in Experiment 2. The English sentences with an imbalanced ratio were meant to be comparable to the Arabic sentences in Experiment 3. Given the nature of English language structure, it was impossible to construct sentences with fewer consonants then vowels, although this would be the ideal design. Nonetheless, as in the case of the Arabic sentences, there were two stimulus presentation conditions, CO and VO, which should sufficiently answer the question under investigation.

In constructing the sentences, I made several decisions regarding segments and words. First, diphthongs were treated as a single vowel, following previous studies such as Cole et al. (1996) and Fogerty et al. (2012). Second, the English rhotic /ɹ/ was treated differently from the Arabic trill /r/. Following Fogerty et al. (2012) and others, I treated /ɹ/ preceded by a vowel as one single rhotacized vowel. Other design aspects that were considered in the Arabic stimuli, such as the inclusion of as many consonants and vowels as possible, the diversity of speech parts, and the diversity of sentence topics were also taken into consideration here. Examples (4) and (5) below show example stimuli with balanced and imbalanced ratios, respectively.

4)	Nobody	saw	you	go	to	Chicago
	/noʊbədi/	/sɔ/	/ju/	/ɡoʊ/	/tu/	/ʃəkɑgoʊ/
	$c^1v_1c^2v_2c^3v_3$	c^4v_4	c^5v_5	c^6v_6	c^7v_7	$c^8v_8c^9v_9c^{10}v_{10}$
5)	Sam	skipped	the	last	three	steps
	/sæm /	/skɪpt/	/ðə/	/læst/	/θri/	/stɛpz/
	$c^1v_1c^2$	$c^3c^4v_2c^5c^6$	c^7v_3	$c^8v_4c^9c^{10}$	$c^{11}c^{12}v_5$	$c^{13}c^{14}v_6c^{15}c^{16}$

4.2.1.4 Judgements. The sentences were sent to native speakers for judgment ratings following the same criteria as in the previous experiments (see Appendix I2). The final version of the stimuli passed the criterion with a rate of 6.11 in written form and 6.20 in spoken form.

4.2.1.5 Recording and stimulus preparation. I followed the same procedure as in Experiment 4, and the design was similar to that of Experiment 4.

4.2.1.6. Tasks. I assigned the same task and instructions as in Experiment 3.

4.2.1.7 Participants. The same participants from Experiment 4 participated in this experiment. The participant who was eliminated in Experiment 4 was also eliminated in this experiment for the same reason.

4.2.1.8 Procedure. I followed the same procedure as in Experiment 4. One additional instruction was that participants were encouraged to report any word they heard out of a presented sentence even if the word as just a definite article (e.g., *the*). The stimuli presentations were counterbalanced. Half of the participants started with imbalanced sentences and half started with balanced sentences. Sentences alternated between the CO condition and the VO condition, but no sentence was presented twice to any participant.

4.2.1.9 Data scoring and analysis. I will use the same scoring and analysis method as in Experiment 3. The number of words reported by the participant for each sentence were converted into percentages, the percentages were converted into angular scores, and the angular scores were then transformed into RAUs.

4.2.2 Results

The overall results showed that speech intelligibility was better in the in the CO condition (47.10% and 47.26 in RAU) than in the VO condition (37.99% and 37.07 in RAU) and in the balanced S-Ratio condition (49.98% and 49.20) than in the imbalanced S-Ratio condition (36.28% and 36.25 in RAU). Figure 14 below compares the results from both conditions. The results also showed that when S-Ratio was balanced, speech intelligibility was higher in the VO condition (62.96% and 62.45 in

RAU) than in the CO condition (37.51% and 36.55 in RAU). When S-Ratio was imbalanced, speech intelligibility was higher in the CO condition (55.49% and 56.61 in RAU) than in the VO condition (16.85% and 16.85 in RAU). Figure 15 below illustrates this set of findings.

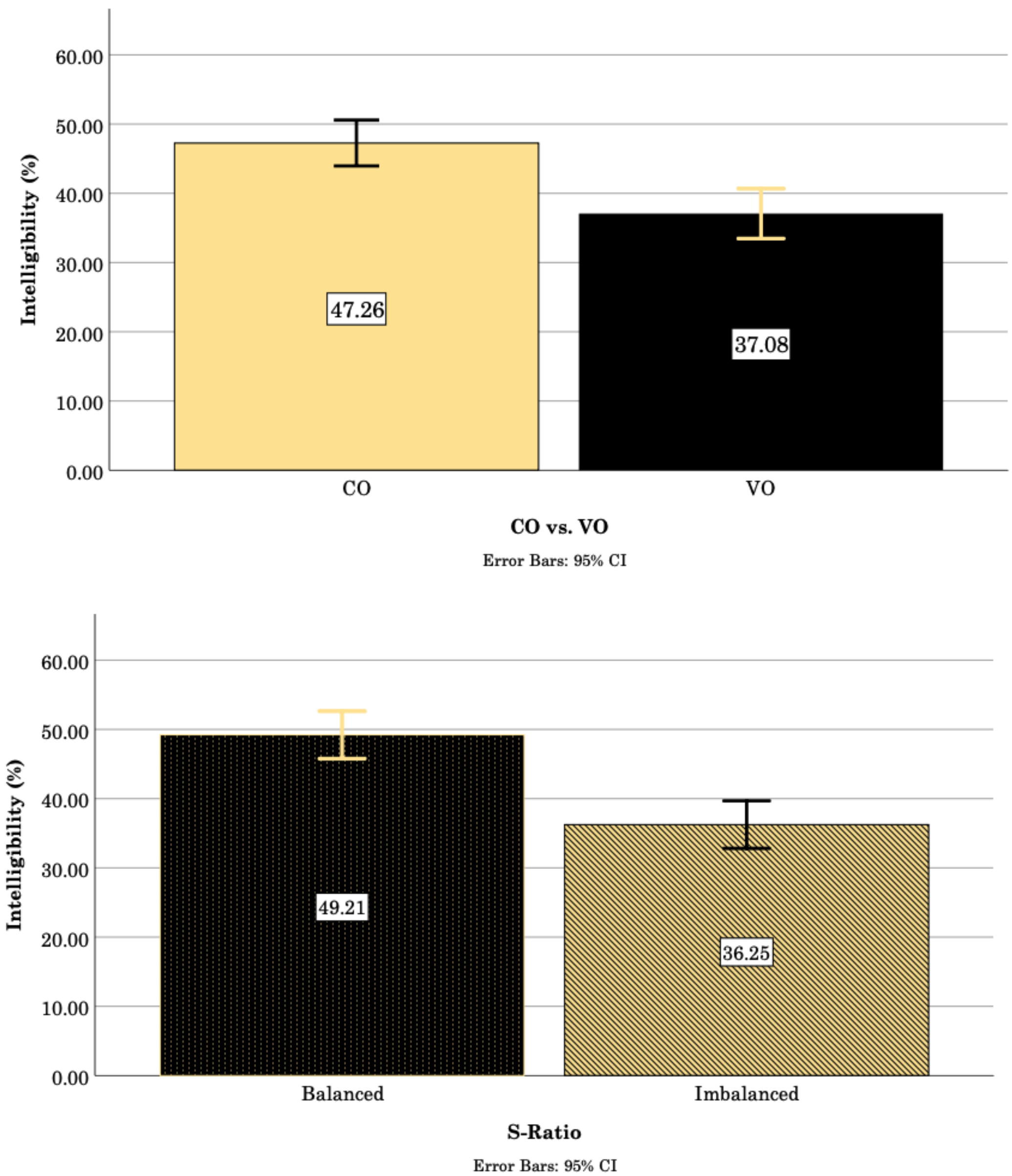


Figure 14. Intelligibility (%) in the C0 vs. VO vs. S-Ratio in Experiment 5

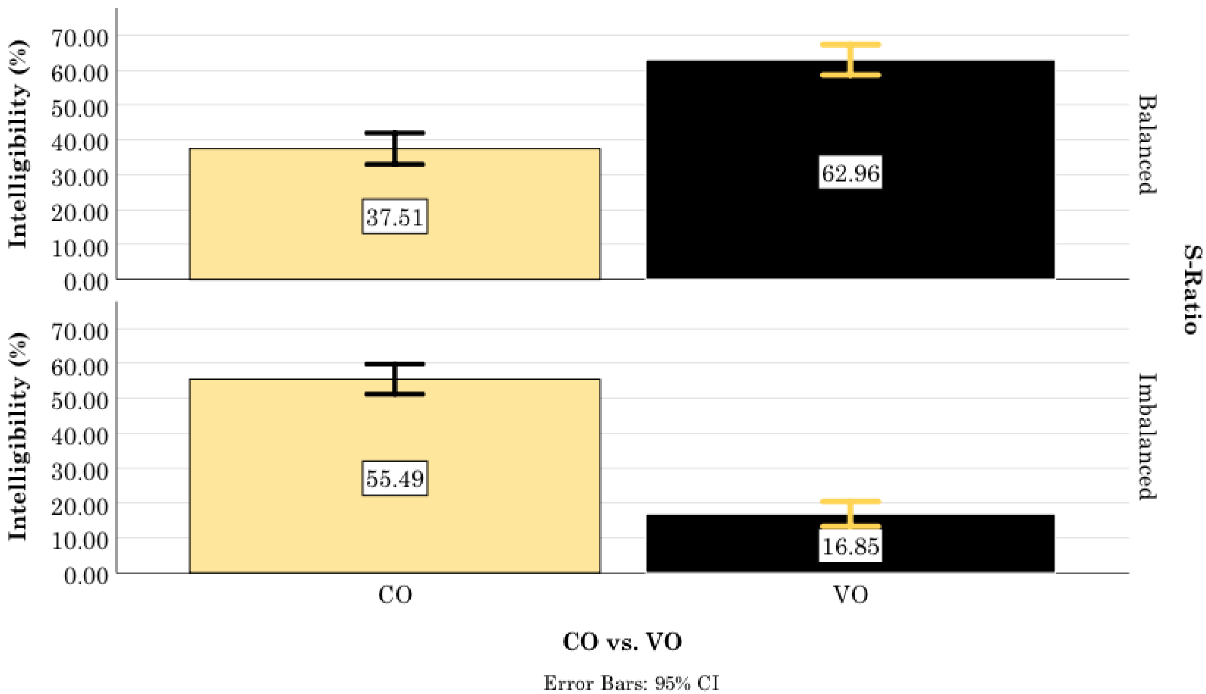


Figure 15. Intelligibility (%) broken down by C0 vs. VO and S-Ratio in Experiment 5

To better understand the results, examine Figure 16 below portraying the mean, median, range, distribution, and frequencies of intelligible words per sentence. The boxplot (Figure 17) shows that, when S-Ratio is balanced, the number of intelligible words ranges between zero and seven words both in the CO and VO conditions but the median (as well as the M) is higher in the VO condition (*MED*= 4, and *M*= 3.76) than in the CO condition (*MED*= 2, and *M*= 2.28). In contrast, when the S-Ratio is imbalanced, the number of intelligible words ranges between zero and six in the CO condition and between zero and two in the VO condition, and the median (as well as the M) is higher in the CO condition (*MED*= 3, and *M*= 3) than in the VO condition (*MED*= 0, and *M*= 0.96).

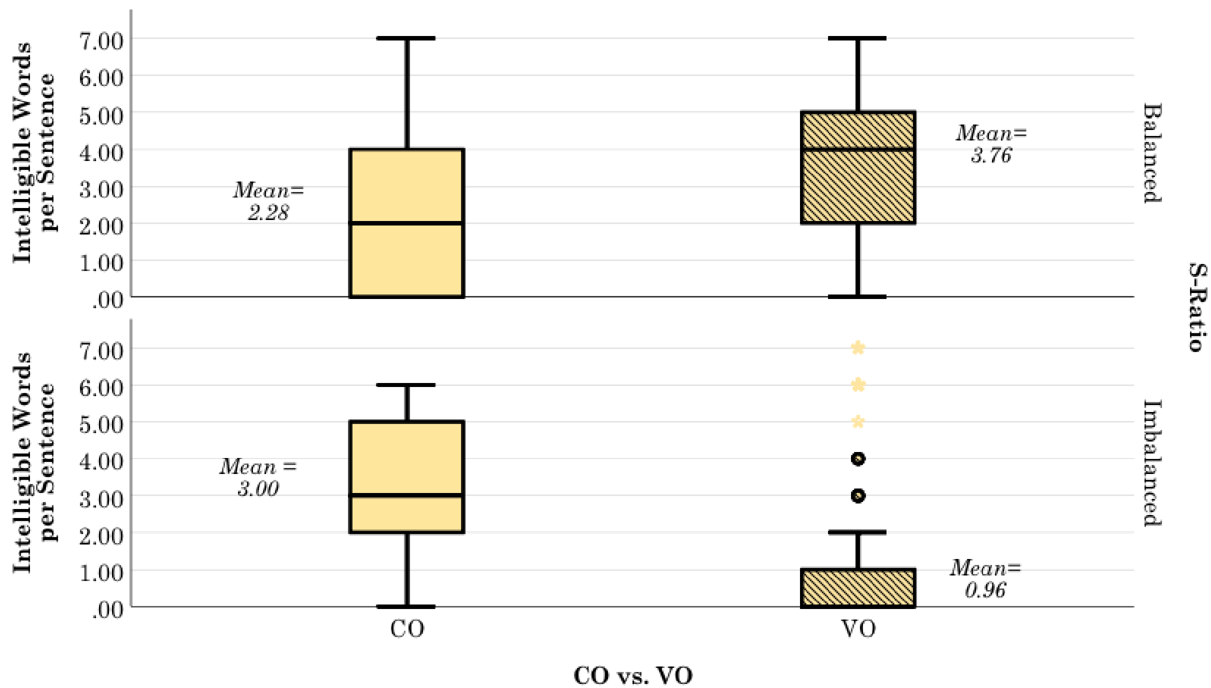


Figure 16. Intelligibility (# of words) broken down by CO vs. VO and S-Ratio in Experiment 5

The dot-plot shows the frequencies of the intelligible words per sentence (see Appendices J2 and K2 for frequencies of RAUs and percentages). We can see that the responses in the CO condition when the S-Ratio is balanced exhibits a similar pattern to that in the VO condition when the S-Ratio is imbalanced. That is, the figure shows many answers with zero intelligible words (*Mode*= 0), which may lead to zero-inflation in these two specific cases. The other two conditions also show a similar pattern that may lead to kurtosis after transformation to RAU. Therefore, when performing a quick exploration of the percentage frequencies (as well as RAU frequencies, as this is the unit used for statistical analysis), we find that the sentences were completely unintelligible 161 times in the VO condition when the S-Ratio was imbalanced and 72 times in the CO condition when the S-Ratio was balanced. Note that, as in the case

of Arabic, some English sentences tended to be more intelligible than the other sentences in both the CO and VO conditions, a side point that will be revisited in the last chapter.

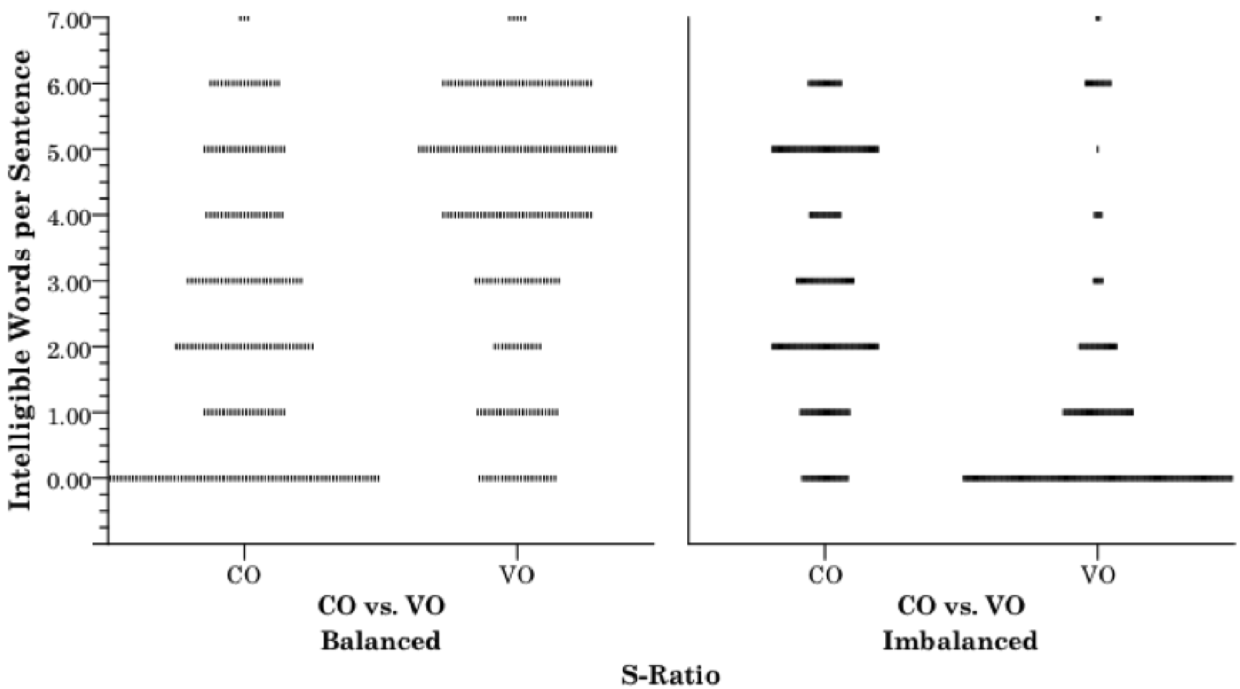


Figure 17. Dot-plot illustration for the frequencies of words per sentence in Experiment 5

This distribution alerted me to the necessity of checking normality before running statistical analyses. I ran a normality test for both independent factors, S-Ratio and CO vs. VO, and the result turned out to be cautious. The normality test revealed a degree of normality, p= 0.01, which again alerted me to further investigate the kurtosis and skewedness of the data. The output revealed a very marginal degree of skewedness ranging between +0.102 and +0.582, and a minor form of kurtosis (platykurtic) ranging between -1.001 and -1.256 in both conditions (see Appendices L2, M2, and N2 for details and detrended Q-Q plots for residuals). There was no need to

perform Mauchly's for the main effect of each factor or apply any correction for the sphericity assumption because there were only two levels per condition. Yet, one remaining concern is the fact that the continuous variable here (RAU) was derived from count data, as was the case in Experiments 2 and 3. However, it was believed that transforming the data from number of words, to percentages, and finally to RAUs was effective in turning the scores into a continuous type. This concern will be revisited and resolved later nonetheless.

After ensuring that it was the appropriate test for the data under investigation, I performed a repeated-measures ANOVA. Table 32 below summarizes the results. The ANOVA test showed evidence of a significant main effect for the CO vs. VO factor at 0.001, $F(1, 19) = 97.359$, $p= 0.001$, $\omega^2 = 0.83$. The effect size was large as shown by the partial eta squared, $\omega^2 = 0.83 > 0.14$. The test also revealed a statistically significant main effect for the S-Ratio factor, $F(1, 19) = 41.304$, $p= 0.001$, $\omega^2 = 0.68$. The effect size was also large (but less than that in CO vs. VO), 0.68> 0.14. The output also indicated a statistically significant interaction effect between the two factors, $F(1, 19) = 325.570$, $p= 0.001$, $\omega^2 = 0.95$. The eta squared value was extremely large, 0.95> 0.14, which indicates that the interaction effect size is truly strong.

Table 32. Summary of repeated-measures ANOVA in Experiment 5

Source		Type III Sum of Squares	*df*	Mean Square	*F*	Sig.	Partial Eta Squared
CO-VO	Sphericity Assumed	63552.803	1	63552.803	97.359	**.000**	.837
	Greenhouse-Geisser	63552.803	1.000	63552.803	97.359	**.000**	.837
	Huynh-Feldt	63552.803	1.000	63552.803	97.359	**.000**	.837
	Lower-bound	63552.803	1.000	63552.803	97.359	**.000**	.837
Error (COVO)	Sphericity Assumed	12402.553	19	652.766			
	Greenhouse-Geisser	12402.553	19.000	652.766			
	Huynh-Feldt	12402.553	19.000	652.766			
	Lower-bound	12402.553	19.000	652.766			
S-Ratio	Sphericity Assumed	24327.765	1	24327.765	41.304	**.000**	.685
	Greenhouse-Geisser	24327.765	1.000	24327.765	41.304	**.000**	.685
	Huynh-Feldt	24327.765	1.000	24327.765	41.304	**.000**	.685
	Lower-bound	24327.765	1.000	24327.765	41.304	**.000**	.685
Error (S-Ratio)	Sphericity Assumed	11190.811	19	588.990			
	Greenhouse-Geisser	11190.811	19.000	588.990			
	Huynh-Feldt	11190.811	19.000	588.990			
	Lower-bound	11190.811	19.000	588.990			
COVO x S-Ratio	Sphericity Assumed	302996.875	1	302996.875	325.570	**.000**	.945
	Greenhouse-Geisser	302996.875	1.000	302996.875	325.570	**.000**	.945
	Huynh-Feldt	302996.875	1.000	302996.875	325.570	**.000**	.945
	Lower-bound	302996.875	1.000	302996.875	325.570	**.000**	.945
Error (COVO x S-Ratio)	Sphericity Assumed	17682.645	19	930.666			
	Greenhouse-Geisser	17682.645	19.000	930.666			
	Huynh-Feldt	17682.645	19.000	930.666			
	Lower-bound	17682.645	19.000	930.666			

The results clearly showed that both CO vs. VO and S-Ratio were critical factors in speech intelligibility and had very strong interactions. The only remaining concern regards the use of RAUs as a linear dependent factor although they were originally transformed from count data (number of words). This was exactly the same issue both in Experiment 2 and Experiment 3. The solution was to run another test, using the number of words as a dependent factor, and the same solution was implemented here. Using GEE, I performed a repeated-measures negative binomial regression analysis (see Appendix O2 for the model details), following the same procedure in Experiments 2 and 3. Table 33 below summarizes the results (see Appendix P2 for full table).

Table 33. Summary of the tests of model effects in Experiment 5

Source	Wald χ^2	*df*	Sig.
(Intercept)	230.428	1	**.000**
CO vs. VO	43.687	1	**.000**
S-Ratio	125.404	1	**.000**
CO vs. VO x S-Ratio	287.324	1	**.000**

Dependent variable: Intelligibility (number of words)

Model: (Intercept), CO vs. VO, S-Ratio, and CO vs. VO x S-Ratio

The model showed that both CO vs. VO and S-Ratio statistically significantly predicted intelligibility, Wald $\chi^2(1) = 230.428$, $p= 0.001$, and Wald $\chi^2(1) = 125.404$, $p= 0.001$, respectively. The model also showed that the interaction between the two factors statistically predicted intelligibility as well, Wald $\chi^2(1) = 287.324$, $p= 0.001$. Table 34 provides the individual parameters and more about the independent factors. The table shows that speech in the CO category was significantly predicted to have a higher *log* count 3.102 times (≈ 75.62%) more than speech in the VO category, and that speech in the balanced category was significantly predicted to have a higher *log* count at 3.901times (≈ 79.59%) more than speech in the imbalanced category, b= 1.132, *SE*= 0.0854, 95% CI = [0.964, 1.300], Wald $\chi^2(1) = 173.662$, $p= 0.001$, Exp(B)= 3.102 (≈ 75.62%), 95% CI = [2.621, 3.671], and b= 1.361, *SE*= 0865, 95% CI = [1.192, 1.531], Wald $\chi^2(1) = 247.784$, $p= 0.001$, Exp(B)= 3.901 (≈ 79.59%), 95% CI = [3.293, 4.621], respectively. The model also significantly predicted that the likelihood of speech in the CO category having a higher *log* count depended on being in the balanced category as few as only 0.196 times (≈ 16.38%) more than in the corresponding category as shown by the negative value of b, b= -1.632, *SE* = 0.0963, 95% CI = [-1.821, -1.443], Wald $\chi^2(1) = 287.324$, $p= 0.001$, Exp(B)= 0.196 (≈ 16.38%), 95% CI = [.162, 236]. In terms of intelligibility, this means that speech in the CO condition was not dependent on being associated with the balanced condition to yield higher intelligibility; speech intelligibility in the CO conditions was higher than in the VO conditions regardless of the S-Ratio.

To recapitulate, both statistical tests showed that speech intelligibility is affected by the type of segmental information (consonantal vs. vocalic) presented to the listener and by the S-Ratio (balanced vs. imbalanced). Higher speech intelligibility in the CO condition at sentence level was a novel finding that was not observed in previous studies, as will be discussed below.

Table 34. Summary of the GEE parameters in Experiment 5

Parameter	B	Std. Error	95% Wald CI		Hypothesis Test			Exp(B)	95% Wald CI for Exp(B)	
			Lower	Upper	Wald χ^2	df	Sig.		Lower	Upper
(Intercept)	-.035	.0854	-.202	.133	.165	1	.685	.966	.817	1.142
[CO vs. VO=1[a]]	1.132	.0859	.964	1.300	173.662	1	**.000**	3.102	2.621	3.671
[CO vs. VO=0]	0[b]	.	.	.	.	.	**.**	1	.	.
[S-Ratio=1]	1.361	.0865	1.192	1.531	247.784	1	**.000**	3.901	3.293	4.621
[S-Ratio=0]	0[b]	.	.	.	.	.	**.**	1	.	.
[CO vs. VO=1] x [S-Ratio=1]	-1.632	.0963	-1.821	-1.443	287.324	1	**.000**	.196	.162	.236
(Negative Binomial)	.325									

Dependent variable: Number of Words per Sentence

Model: (Intercept), CO vs. VO, S-Ratio, and CO vs. VO x S-Ratio

a. CO was coded as 1, VO as 0, Balanced as 1, and Imbalanced as 0.

b. Set to zero because this parameter is redundant.

4.2.3 Discussion and Conclusion

Experiment 5 was undertaken to examine the effect of S-Ratio on speech intelligibility at the sentence level in English. The results showed that speech intelligibility was better in the CO condition than in the VO condition when the S-Ratio was imbalanced (more consonants). This is an unprecedented finding that diverged from all previous studies that have showed that vocalic information is more contributory than consonantal information at the sentence level. The results from the balanced ratio condition, however, agree with previous studies. These two outcomes indicate that both segment type and S-Ratio interact and affect speech intelligibility. The discussion of S-Ratio should be accompanied with the discussion of segmental contributions.

4.2.3.1 Segmental contributions. At the first glance at the visual representation (Figure 15) one may suppose that, if the S-Ratio is balanced, vowels contribute to speech intelligibility more than consonants do, but if the S-Ratio is imbalanced, then consonants contribute to speech intelligibility more than vowels do. The interaction between the two factors made it difficult to conclude which segment is more contributory than the other. This is similar to the case of initial segment in which vowels contributed more than consonants in one condition, but consonants contributed more than vowels in another condition. The solution was to group the results into one category and look at the grand mean instead. Our interpretation of the overall results is that vowels and consonants appear to contribute equally to speech intelligibility in English. However, because vowels inherently carry some additional communicative

information at the sentence level, they have an additional advantage that makes them appear more contributory than consonants. With a balanced ratio, speech intelligibility should be approximately even (50.00%) in the CO and VO conditions, but here the rate dropped 12.49% in the CO condition and increased 12.96% in the VO condition below and above 50.00%, respectively. The increase in the VO can be interpretable as due to supralinguistic cues, rather than linguistic properties, that vowels enjoy. As Fogerty et al. (2009; 2012) described it, "the spectral differences alone do not explain the observed greater vowel contributions in sentences" (Fogerty et al., 2012, p. 1676). One evidence for this view can be sought and found in Chen et al. (2013) who demonstrated a larger intelligibility advantage for vowels over consonants in Mandarin Chinese (3:1) than in English (2:1). They attributed this difference to the fact that vowels in Mandarin are carriers of tonal information that scaffolds sentence intelligibility in the VO condition, as such information is important for word recognition. Languages will differ in this respect, creating additional difficulty for experimental design unless we can manage to structure some sentences with the least contextual information that vowels carry out. For instance, we may attempt constructing sentences out of individual words recorded in isolation, as will be discussed in Chapter 5.

Now, consider the second half of the results regarding intelligibility in the imbalanced ratio. If vowels were more important for speech intelligibility in English, we would expect a very low speech intelligibility in the CO condition for the imbalanced

ratio. This did not occur. Instead, with only a 10-segment difference (15C-5V), intelligibility rate was higher than 50.00% in the CO condition, but remarkably poor in the VO condition. This condition is more representative of English language structures than is the balanced condition, because English sentences usually have more consonants than vowels (Fogerty et al. 2009), although the ratio here (3:1) is more incommensurate than the ratio (2:1) in Fogerty et al. (2009). The question poses itself again: does the pattern for the imbalanced ratio mean that consonants contribute more information to the intelligibility of sentences in English? The answer, I believe, is again *no*. We were successful at manipulating the ratio to make the segment that was always reported as contributing more to speech intelligibility appear to contribute less. It is evident that we can construct English sentences in which vowels contribute more and others in which consonants contribute more, which suggests that both are important and contribute equal information to intelligibility. However, if one type had a higher privilege, then it would maintain its privilege regardless of the S-Ratio. This is utterly different from the scenario in nonconcatenative languages where one type is given a degree of superiority by the language system. In that case, we failed to make vowels more important in all situations regardless of the non-spectral information they carry at the sentence level.

In summary, vowels and consonants appear to equally play a key role in English intelligibility, but for other factors such as S-Ratios, one segment may appear as being more important than its counterpart in one context or another. The S-Ratio

played an important role in demonstrating this equality, and although it has been included in the discussion above, I will touch upon it again below.

4.2.3.2 Segmental ratios. S-Ratio has previously been considered and examined as a potential factor in the literature. However, the type of ratio tested in previous studies was different from the one manipulated and examined here. Previous studies tested the CV-Ratio of segments within a language overall. Cutler et al. (2000) raised and tackled the question as to whether CV-Ratio would impact segmental contributions to intelligibility. They manipulated the CV-Ratio, using Spanish, which has an imbalanced ratio with more consonants than vowels, and Dutch which has a balanced ratio with a relatively equal number of consonants and vowels, but found no significant effect for this factor. In the current experiment, S-Ratio in English sentences (as well as in English words in previous experiments) revealed a significantly new pattern at the sentence level, which was the same domain (i.e., sentences) in which vowels exhibited privilege in previous studies such as Cole et al. (1996) and Fogerty et al. (2007, 2009, and 2012). The imbalanced ratio did not only quantify the acoustic vocalic information, but also limited the impact of sentential context on speech intelligibility. This collectively resulted in poor performance in the VO condition for the imbalanced ratio.

The vocalic information and consonantal information in the original sentences for the imbalanced ratio constituted 33.33% and 66.66% of the speech content, respectively. Once a third of speech content is removed (as in the CO condition), we would expect the speech intelligibility rate to be nearly twice the speech intelligibility rate

once two thirds are removed (as in the VO condition). For example, assume that speech intelligibility in the VO condition was an average of two words out of six. Hence, with everything else being equal, speech intelligibility would be expected to be about four words in the CO condition. In our results, the intelligibility proportion in the VO condition was 16.85%, and based on this rate, the speech intelligibility proportion in the CO condition would be expected to approximately 33.70%, but the obtained rate was much higher (ratio= 3.29:1). This means that consonantal segments truly provided rich and useful information that helped listeners understand speech to a degree similar to that in the VO condition for the balanced ratio.

An alternative explanation would be based on the low rate in the imbalanced-ratio VO condition rather than the CO condition. The intelligibility rate was 55.49% in the CO condition, and it would be expected to be approximately 27.74% in the VO condition if vowels and consonants are equal contributors, but it was instead less than that expected value. It seems that speech in this condition (i.e., the VO condition) was probably extremely interrupted with a third of the speech replaced by silence. This explanation would also account for the slight unclear drop below 50.00% in the balanced-ratio CO condition. There were too many interruptions for the speech stream, as half of the content was silence-replaced. Interruption must have occurred in the VO condition too, but we have no means to quantify its impact on intelligibility in the presence of the sentential context that affects vowels differently from consonants. It is possible that silence interruption occurring to consonants is different from interruption occurring to vowels.

To summarize, the S-Ratio had a differential impact on vowels and consonants. A balanced ratio resulted in a higher intelligibility rate for the vocalic information, whereas an imbalanced ratio resulted in a higher intelligibility rate for the consonantal information. When grouping the outcomes from both S-Ratio types, consonants appear to play a greater role than vowels do. This indicates that segmental contributions are divided between vowels and consonants in English, but some intervening factors manipulate their contributions to make each appear more contributory in a specific context.

⇐ Chapter Four

⇒ References

CHAPTER FIVE:

GENERAL DISCUSSIONS AND CONCLUSIONS

5.1 An Overview

This dissertation contrasted segmental contributions and S-Ratios between languages with nonconcatenative morphological systems (represented by Arabic) and languages with concatenative systems (represented by English). In doing so, a series of five experiments – three on Arabic and two on English – were conducted to examine the effect of major factors on speech intelligibility: nonconcatenative vs. concatenative language, word vs. sentence level, balanced vs. imbalanced ratio, and vocalic vs. consonantal information. The design of the experiments also considered additional factors such as initial vowel vs. initial consonant, initial position vs. final position, and noun vs. verb, while only the initial segment turned out to be an important factor in speech intelligibility at the word level.

It was predicted (and confirmed) that, due to substantial differences between nonconcatenative and concatenative languages with regard to their morphological architecture, segments contribute differently to speech intelligibility in Arabic compared to in English. To be more precise, speech intelligibility in Arabic benefits more from consonantal information than from vocalic information, and S-Ratios affect the contributions of consonants as the primary carrier of speech information more than they affect vowels. Speech intelligibility in English equally benefited from both segment types in these experiments, with S-Ratios having an impact on both types.

Five experiments were conducted. In the first experiment, the Arabic participants were presented with Arabic lexical items with different S-Ratios (balanced vs. imbalanced) while either vocalic information or consonantal information was uniquely present. The outcome illustrated a high importance of consonantal information for Arabic speech to be intelligible. In the second and third experiments, the participants were presented with complete sentences with either balanced or imbalanced S-Ratios, while either vowels or consonants were retained in the stimuli as in the first experiment. The results again showed that consonantal information remained the major contributing segment to speech intelligibility in Arabic, although more vocalic information improved speech intelligibility to a degree. Taken together, this provided evidence that the Arabic morphological system, which mainly assigns lexical and semantic information to consonants and morphosyntactic information to vowels, plays a fundamental role in speech intelligibility for Arabic speakers. Such a conclusion diverges from previous research that has emphasized the role of vocalic information in speech intelligibility at the sentence level in English, and from previous research that has stressed the role of vocalic information to speech intelligibility at the word level in Chinese. However, it concurs with a claim expressed often implicitly, and rarely explicitly, in the literature that segmental contributions are language-specific.

In the fourth experiment, which was designed to resemble the first experiment, the English speakers were presented with English lexical items that started with either a consonant or a vowel and had either a balanced or imbalanced ratio, using

the same silence-replacement paradigm. The findings uncovered a new and important factor that was overlooked in the literature, namely initial segment, which showed that vowels contribute to speech intelligibility more than consonants do once the initial segment belongs to the same category (vowels), and vice versa. In the fifth experiment, which was designed to resemble the second and third experiments in Arabic, English participants were presented with segment- balanced vs. imbalanced sentences in which only consonantal or vocalic information was preserved. The balanced sentences triggered a higher speech intelligibility when only vocalic information was presented, whereas imbalanced sentences with more consonants resulted in a higher speech intelligibility when only vocalic information was presented to the participants. Combined together, the results from the English experiments provide evidence that vowels and consonants play roughly the same role in speech intelligibility in concatenative languages, a conclusion that reconciles disputing outcomes reported in previous research.

This manifestly different role of segments in Arabic vs. English under similar conditions is attributed to the difference between the morphological system that each language employs. That is, as Arabic assigns different roles to vowels and consonants, consonants, which carry lexical information, act as a primary source for Arabic speakers to process and comprehend speech. The absence of such a distinction in English preserved equality in segmental contributions.

5.2 General Discussion

Two major topics are brought up for discussion here: segmental contributions in nonconcatenative vs. concatenative languages and S-Ratios in concatenative vs. nonconcatenative languages. The discussion will also cover some important sub-topics such as the effect of isolated words vs. complete sentences, and the effects of initial segments. In addition, the concept of bottom-up vs. top-down processing will be an important aspect of the discussion later on.

Research has established two preeminent modes of speech processing referred to as bottom-up and top-down. In bottom-up processing, computation of segmental information is the fundamental operation that occurs first. This approach has been supported by many advocates in speech science and psycholinguistics (see e.g., Cutler, Mehler, Norris, & Segui, 1987; Norris, McQueen, & Cutler, 2000). Some studies even went beyond arguing for bottom-up processing to stress that, since speech processing is monodirectional in a bottom-up manner, top-down processing could *hinder* speech intelligibility (e.g., Norris et al. 2000). Such a thesis states that the flow of information from segments to words (or even sentences) is always necessary for speech intelligibility, but a backward feedback from words to segments is unnecessary or probably implausible.

In a top-down model, on the other hand, higher levels of information (i.e., non-segmental information) is processed to comprehend speech, even with limited segmental information (see e.g., Connine & Clifton, 1987; Marslen-Wilson, 1984; McClelland, & Elman, 1986; Samuel, 1996). For example, if the listener believes she has

recognized the word “thrift”, this higher-level information influences her perception of lower-level phonemes, biasing her to believe she has perceived the segments /θ/, /ɹ/, /ɪ/, /f/, and /t/. Full sentences provide the canonical example of top-down information, because they can provide sufficiently strong contexts for a listener to essentially “re-store” *entire words* without access to bottom-up information from speech. For example, provided with the sentential context “The touch of soft fur and the sound of a meow made me realize that a ____ had entered the room”, English listeners will invariably hear “cat”.

Models of word recognition include the Logogen Model (Morton, 1969), the Schema Model (Reumelhart, 1977) the Cohort Model (Marslen-Wilson, 1978), the RACE model (Cutler & Norris, 1979), TRACE (McClelland & Elman, 1986), the Fuzzy Logical Model of Perception (FLMP) (Massaro, 1989; Massaro & Cohen 1991), Shortlist (Norris, 1994), the Distributed Cohort Model (Gaskell & Marslen–Wilson, 1999), and the Merge Model (Norris et al. 2000). These models differ primarily in terms of how they implement top-down vs. bottom-up approaches. The dispute between the two approaches is centered around irreconcilable evidence from different experimental tasks and techniques (Norris et al., 2000) such as phonemic categorization (McQueen, Norris, & Cutler, 1999) and phonemic restoration (Samuel, 1996). Defined as the phenomenon where some phonemes absent from speech can be *restored* and believed to be heard by the listener, phonemic restoration is in fact similar

to the technique implemented in the current study. Our participants used one segment type (consonants or vowels) to restore the other type and then comprehend the intended speech.

A number of previous studies attempted to evaluate the amount of bottom-up vs. top-down information needed for speech intelligibility (see e.g., Samuel, 1981, 1996, and 2001), using words and sentences. That is, such studies tried to examine how much intelligibility occurs as a function of segmental information vs. non-segmental information, such as sentential context. In general, the conclusion from this series of studies (see Samuel, 1996) was that the top-down effect is real but fragile, and the necessity of an interactive account for top-down and bottom-up modelling may be questionable. Nevertheless, other studies such as Fogerty et al. (2012) and Chen et al. (2013) suggest that the effect of top-down processing on speech intelligibility is strong.

In the current study, we relate this concept (i.e., top-down vs. bottom-up processing) to the findings on segmental contributions to speech intelligibility, but before we do so, we should first discuss the overarching findings from the two previous chapters (Chapter 3 and Chapter 4).

5.2.1 Segmental Contributions in Nonconcatenative vs. Concatenative Languages

Over the years of research recorded in the literature, a handful of studies suggested that consonants play a prominent role in speech intelligibility. For instance, Owens et al. (1968) stated that "it [had] become almost a commonplace statement in

intelligibility testing that most of the information in speech is carried by the consonant sounds" (p. 648). Kewley-Port et al. (2007) also indicated that "the common wisdom in audiology is that the most important segments for word recognition are the consonants, not the vowels" (p. 2370). This belief can be traced and dated back to 1929 in a study by Fletcher (1929)[17] and Dudley (1940) who focused on errors in phoneme identification for monosyllabic words and tested the carrier nature of speech for their purpose (a transmission system for speech). Most of the early work on this topic (as well as some later studies such as Owren & Cardillo [2006]) used nonsense words or otherwise used words mainly with CV or CVC structures. However, Cole et al. (1996) challenged this commonplace view, using complete sentences, and reported that vowels contribute more than consonants to speech intelligibility. This new finding sparked a debate and led to a consecutive set of studies investigating segmental contributions to speech intelligibility (from van Ooijen [1996] to Chen et al. [2015] to the current study).

Subsequent studies attempted to test this outcome further and probe additional factors influencing segmental contributions to intelligibility. Cutler et al. (2000) thought that language CV-Ratio (recall that CV-Ratio is the number of vowels vs. consonants in the inventory of a given language) could influence the contribution of a single segment in one language or another. Their findings showed that consonants played a greater role than vowels in speech intelligibility at the word level, but they did not support the possibility that the CV-Ratio effect was a hidden factor in

[17] For an overview on Fletcher's contributions to speech perception and communication acoustics see Allen (1996).

previous studies such as van Ooijen (1996). Segment margins, especially for vowels, were also examined as a factor in several studies such as Cole et al. (1996) and Owren and Cardillo (2006). The results of their studies did not uphold this hypothesis, however. On the other hand, the additional information that vowels are thought to convey at the sentence level also attributes to them an important role with respect to speech intelligibility. Fogerty et al. (2012) compared segmental contributions to intelligibility in individual words vs. complete sentences, devoting a proportion of their study to analyze what contextual benefits were provided to vowels vs. consonants in sentences as compared to individual words. Their analysis revealed a difference between the contextual benefits associated with vowels and consonants; consonants did not gain any contextual benefits at the sentence level while vowels by comparison did. Hence, they concluded that "spectral differences alone do not explain the observed greater vowel contributions in sentences" (p. 1676).

Fogerty et al.'s (2012) conclusion could be interpreted as an indication that segmental contributions are not only tied to the type of acoustic properties of vowels vs. consonants; rather, one segment type may contribute richer information than the other depending on its context, such as within a sentence. This allows listeners to use such non-acoustic information for top-down processing, which, along with bottom-up processing, enhances speech intelligibility. It follows that a language that employs one of the two classes of segment types in a particular way – for example, for morphological reasons, as in Arabic – is likely to behave differently. This forms the founda-

tion of the hypothesis of this dissertation regarding an opposition between concatenative and nonconcatenative languages. Nonconcatenative languages assign a substantial role to consonants in their lexical architectures – namely, consonants carry out lexical and semantic information while vowels carry out structural information. English, as we have seen, does not. In support of this hypothesis, our five experiments show consistent differences between the results for these two languages. The findings, regardless of the experimental design manipulations with respect to word vs. sentence, balanced vs. imbalanced, noun vs. verb, and initial vs. final, exhibit abundant evidence that consonants in Arabic, but not in English, make stronger contributions to speech intelligibility than vowels do.

We should therefore be in a secure position to posit that segmental contributions to speech intelligibility in both concatenative and nonconcatenative languages can be attributed to both acoustic (i.e., spectral) and non-acoustic information, and that such types of information are used for top-down vs. bottom-up processing differently in Arabic vs. English. In concatenative languages, such additional information is provided by the nature of speech production at the sentence level which allows vowels, but not consonants, to convey the major portion of contextual information that helps listeners to process speech using some sentence-based cues. In other words, instead of phoneme-by-phoneme (i.e., bottom-up) processing which usually occurs at the word level, speech intelligibility at the sentence level is likely to benefit from top-down processing. This hypothesis, which seems to have originated from previous studies such as Rumelhart's (1977) and Samuel's (1981, 1996, and 2001) works, was

alluded to by Owren and Cardillo (2006) while comparing their results with Cole et al.'s (1996) and used explicitly by Kewley-Port et al. (2007) and Fogerty et al. (2012) as a justification for the discrepancies between different studies including their own. In contrast, the additional non-spectral information in nonconcatenative languages stems from the morphological systems they respectively employ. That is, as many previous studies have shown, lexical items in nonconcatenative languages are mainly built of a string of consonants that get interdigitated with another string of vowels (see e.g., Holes, 2004; McCarthy, 1981; Watson, 2007, among many others). Arabic speakers thus benefit more from consonantal information than from vocalic information to structure the mental dictionary and access words, as a large body of research has discussed (see e.g., Abu-Rabia, 1997; Bentin, and Feldman 1990; Deutsch and Frost 2003; Feldman, Frost, and Pnini 1995; Deutsch, Frost, Pollatsek, and Rayner 2000; Feldman; 2000; Holes, 2004; Minouni, Kehayia, and Jerema, 1998; and Ravid and Shimron, 2003). The current findings are therefore in line with previous findings. The purpose of the current study, however, was to specifically evaluate segmental contributions to speech intelligibility in a nonconcatenative language (Arabic), making it different from previous studies on Arabic and Hebrew. Furthermore, all previous works used complete or partial isolated words or nonwords, and none of them used complete sentences. The use of sentence stimuli in this study revealed that Arabic speakers also benefit from vocalic (and hence structural) information at the

sentence level, but not as much as they do from consonantal information. The following sub-sections discuss segmental contributions in Arabic and English both in words and in sentences.

5.2.1.1 Segmental contributions in Arabic words vs. sentences. As stated above, consonantal information turned out to be superior to vocalic information at both the word and the sentence levels in Arabic, which deviates from previous findings about concatenative languages. Interestingly however, vowels do become more important in sentences than they are in words: in the VO condition, speech intelligibility was notably higher at the sentence level than at the word level. A plausible question to pose is whether or not the structural information carried by vowels plays a role in speech intelligibility in Arabic. The non-spectral information in the VO condition stems from the morphosyntactic information, while the non-spectral information in the CO condition comes from the morphological system assigning semantic information to consonants. On this basis, we might have expected contributions to intelligibility to be similar or equal for both segment types, because vowels carry structural information and consonants carry lexical information.

In the end, speech intelligibility in the VO condition benefited from the structural information available at the sentence level, but this did not sufficiently compensate for the absence of consonantal information. This indicates that the effect of the Arabic lexical information associated with consonants is stronger than the effect of

structural information associated with vowels in Arabic. Put differently, lexical information seems to be more important for overall intelligibility than structural information in Arabic.

Recall that many studies on Arabic (and Hebrew) have shown that the consonants constituting the root form the morphological unit for word access, activation, and recognition (see e.g., Bentin & Feldman, 1990; Feldman, Frost, & Pnini, 1995). Vowels forming the pattern have been considered as a secondary unit in some forms of verbal items (see e.g., Deutsch & Frost, 2003), but nevertheless fail to exhibit an equivalent importance when compared to consonants. Taking all this together, vowels contributed much less than consonants did to speech intelligibility at the word level in Arabic. Put differently, structural information alone was not a reliable unit for speech intelligibility, in agreement with what most previous studies reported about the use of consonants vs. vowels as a lexical access morphemic unit. Such studies used words, partial words, and nonsense words for their experimental stimuli, however none used sentences, potentially because their experimental interest lay in modeling a mental lexicon for Arabic rather than in investigating speech intelligibility.

It is likely that the vocalic information helped the Arabic listeners to identify a word's structure and locate it in the sentence. In taking a closer look at results from individual trials, there were many instances in which the participants successfully recognized the word structure (correct vowels) but failed to recognize the intended words (incorrect consonants). For instance, a verb-subject (VS) sentence in Experiment 2 starting with the two words /taħaddaθa/ "talked" /muħalliluun/ "analysts"

(analysts talked) was reported as /takallama/ “spoke” /muħaddiθuun/ “talkers” (talkers spoke) in the VO condition by several participants. The vowels are identical in both the presented sentence and the reported sentence, but the intended words were missed. The structural information facilitated intelligibility of the word and sentence structures, since the participants were able to identify the first item as being verbal and the second as being nominal. The structural information here includes the verb tense which is carried by the a-a-a pattern, the agreement morpheme which is represented by the last vowel, and pluralization, gender, and nominative case which are all represented by the long vowel /uu/. According to the scoring technique used in the present study, such responses were regarded as incorrect, but they still serve as an indication that there were some other instances in which such structural information helped the participants to recognize the intended lexical item. This is taken as evidence that structural information enhanced speech intelligibility in the VO condition at the sentence level in Arabic. In addition, the extremely low speech intelligibility in the VO at the word level compared to the sentence level was probably affected by the lack of stimuli containing initial vowels. Because of constraints on the Arabic lexicon, there was no counterpart condition in which vowels served as the initial segment.

In summary, the current study has shown that consonantal information contributes to speech intelligibility more than vocalic information does in Arabic both at the word and sentence level, and that other factors did not interrupt the overall es-

tablished pattern. In other words, segmental contributions in Arabic remained consistent across both words and sentences, an observation that did not hold true for English as will be discussed below.

5.2.1.2 Segmental contributions in English words vs. sentences. Apart from some experimental design differences, the experiments conducted on English resemble those conducted on Arabic and address one of the chief issues investigated in this dissertation. Specifically, the English experiments were designed to assess the impact of S-Ratio on speech intelligibility in English. The outcome, however, led to some interesting and important realizations about the initial segment at the word level and S-Ratio at the sentence level. As we saw in Chapter 4, the discussion of segmental contributions cannot be completely disassociated from S-Ratio and initial segment. Nevertheless, the overall results showed that vowels and consonants contribute almost the same amount of information to speech intelligibility in English, a new finding that did not emerge in the literature known so far.

5.2.2 Segmental Ratios in Nonconcatenative vs. Concatenative Languages

S-Ratio on its own is not an independent factor that determines speech intelligibility, but given that it affects the magnitude of consonantal vs. vocalic information available to the listener, it was a useful variable to manipulate. It also intersected with some other variables such as initial segment and word vs. sentence, and showed how segmental contributions are not always static. In general, for Arabic speakers, manipulating the S-Ratios at the word level showed that an imbalanced ratio with

more vowels than consonants did not improve speech intelligibility; however, the reverse did. For English speakers, manipulating S-Ratios at the word level affected both segment types almost on an identical basis. The same effect transmitted identically to the sentence level, creating a similar pattern in speech intelligibility.

5.1.2.1 Segmental ratios in Arabic words vs. sentences. The three experiments on Arabic revealed that the S-Ratio of the stimuli played a role in speech intelligibility but remained overridden by the strong role played by vowels and consonants themselves in Arabic, regardless of whether words or sentences were being examined. This indicates that, unlike in English, the morphological system of Arabic plays a dominant role in determining how much each segment contributes to speech intelligibility. At the word level, a one-consonant difference was sufficient to increase or decrease the likelihood for speech to be intelligible or unintelligible. This was taken as evidence that consonantal information is the primary source for lexical information in Arabic, which concurs with what has been reported in studies such as van Ooijen (1996), Cutler et al. (2000), and Owren and Cardillo (2006). The interpretation presented in the current study is different, however. I attribute this consonantal privilege to the role consonants play in the nonconcatenative morphological system of Arabic, and claim that a lack of consonantal information hampers bottom-up processing. That is, with a 3C-3V ratio, bottom-up processing for Arabic listeners proceeds on the basis of three segments (consonants); with a 2C-3V ratio, however, bottom-up processing proceeds on the basis of only two segments, and suffers as a result.

This interpretation may not be convincing unless we compare it with S-Ratio at the sentence level, because previous results have shown that consonants also play the dominant role in nonconcatenative languages. The answer to this valid objection is simple; consonants remain superior at the sentence level in Arabic regardless of the structural information that vowels provide. Moreover, in sentences in the CO condition, reducing the number of vowels in the target stimulus (i.e., an imbalanced ratio of V to C) resulted in better speech intelligibility than when equal numbers of vowels and consonants were presented (i.e., a balanced V to C ratio). This consonantal superiority can be further supported if collapsing the results from the balanced vs. imbalanced ratios into one outcome maintains the role of consonants intact. In fact, this was exactly the case (see Appendix Q2).

In summary, the pattern of segmental contributions found in Arabic consistently presents consonants as the primary contributing segment to speech intelligibility. S-Ratios amplify this finding; a ratio that imbalances the vocalic vs. consonantal information in favor of consonants increases speech intelligibility. Again, it is believed that this is attributable to the dominant position that consonants occupy in the Arabic morphological system, which makes consonants the fundamental unit for lexical architecture and consequently for speech processing and intelligibility. The scenario in English is radically different, as presented below.

5.1.2.2 Segmental ratios in English words vs. sentences. In English, the importance of vowels vs. consonants depended crucially upon the experimental condition. In some situations, vowels appeared to contribute more than consonants to

speech intelligibility, such as when the initial segment was a vowel in words with balanced ratios. In some other situations, consonants manifested a greater contribution than vowels did, such as when the segments in sentences were balanced.

At the word level, the initial segment played a critical role. Although the experimental design (i.e., the absence of an initial vowel in the imbalanced ratio condition) made it difficult to directly compare the results from the balanced and imbalanced ratios, vowels and consonants played a similar role when the ratio was balanced. The novel finding from the word-based experiment was the way that words with initial vowels inverted the pattern reported in the literature on English, such as in Owren and Cardillo (2006). In the present study, for vowel-initial words, vowels in the VO condition appeared to contribute more than consonants in the CO condition to intelligibility. If vowel-initial words were available in Arabic, we would not expect to see a similar pattern, since the initial vowel alone would not help Arabic speakers to process speech. That is, consonantal information, even if not in an initial position, would still be necessary to facilitate bottom-up processing and speech intelligibility in Arabic.

At the sentence level in English, it was manifest that manipulating S-Ratios modulated the contributions of vowels vs. consonants. This would mean that, unlike in Arabic, collapsing the results from both conditions (balanced vs. imbalanced) should make speech intelligibility rates in the CO and VO conditions relatively approximate. Indeed, this was what the outcome turned out to be (see Appendix Q2).

When screening results for the sentence-based experiments in relation to previous studies on English such as Cole et al. (1996), Kewley-Port et al. (2007), and Fogerty et al. (2012), a few differences become apparent. To lay this out, recall that Cole et al. claimed that in their stimuli "the sentences were selected to have the same number of vowels and consonants" (p. 854) and that "the experiment was designed to [balance] the numbers of occurrences of consonants and vowels to the extent possible within each sentence" (p. 854). If Cole et al. (1996) had a truly balanced ratio, then our results from the balanced ratio here are perfectly in agreement with their results. However, it is more likely that they did not have a balanced ratio, an issue that was fully explained in Chapter 2, simply because they excluded weak sonorants from the class of consonants. This would mean that their stimuli actually had an imbalanced ratio (more consonants than vowels), similar to the imbalanced ratio we investigated in this study. Nevertheless, our results are completely different from those in Cole et al. While Cole et al. had a higher speech intelligibility in the VO condition, we had an overall higher speech intelligibility in the CO condition. This could be interpretable as the result of multiple differences in design. First, in their first experiment, the weak sonorants were available in both the CO and VO conditions. This would mean that the VO condition was not truly a VO condition; rather, some consonants (weak sonorants) were available to the listeners. Second, in their second experiment, the weak sonorants were removed from the CO condition. This would mean that the CO condition was not truly a CO condition; instead, a great portion of consonants were removed from speech, which definitely affected S-Ratios, segmental contributions,

and hence speech intelligibility. Third, in their fourth experiment, in addition to the issue outlined with respect to their second experiment, they also shrank and expanded vowels and consonants (refer to Figure 1), which is a technique that was not used in this study. Therefore, although their experiments and Experiment 5 here can be said to both have an imbalanced ratio, the way segments were classified and substituted is not comparable.

In some subsequent works such as Kewley-Port et al. (2007), Fogerty et al. (2009), and Fogerty et al. (2012), similar findings to Cole et al. (1996) were reported. In these studies, the authors reported that they used the same sentences in all three studies, but it was only Fogerty et al. (2009) in which a description of the S-Ratio was documented. They stated that the average length of the sentences was 11 vowels and 22 consonants. This informs us that they perfectly but not purposely implemented a segment difference (11-segment difference) that was very similar to the one implemented in the current study (10-segment difference), regardless of the total lengths in their study and our study (33 segments vs. 20 segments, respectively). Notwithstanding, our results from the imbalanced ratio diverge from those reported in the three studies in question. In particular, vowels contributed more than consonants in their studies, while consonants contributed more than vowels did in our study, similar to the case of Arabic. One possible explanation is the disparity between the number of vowels in our design (5 vowels) which are less than half those in their study (11 vowels). Such an observation emphasizes the vowel's capability of carrying supraseg-

mental information in English. With as few as five vowels in a sentence, speech intelligibility suffered from the lack of both spectral information necessary for bottom-up processing and suprasegmental information necessary for both bottom-up and top-down processing. When we increased the number of vowels in the sentence-balanced VO condition, participants were able to perform much better than when only five vowels were present to them (in the sentence-imbalanced VO condition).

Consider now the following difference between English and Arabic. Reducing the number of vowels in a sentence severely impacted speech intelligibility in the sentence-imbalanced VO condition for English speakers. This indicates that the sentence-level cues English speakers take advantage of deteriorated in the VO condition. Contrariwise, reducing the number of vowels in Arabic did not cause a huge difference in speech intelligibility; as a matter of fact, it ameliorated speech intelligibility in the CO condition. In other words, it was not surprising to see an increase in speech intelligibility for the English CO condition because we increased the amount of consonantal information (5 segments were increased to 10 segments), but it was surprising to see an increase in speech intelligibility for the Arabic CO condition since the ratio was manipulated by reducing the number of vowels rather than increasing the number of consonants. This is fascinating, because it shows that although Arabic speakers may take advantage of vowels carrying structural information (as well as suprasegmental cues) in the VO condition, they processed and comprehended speech better when it originally had fewer vowels and was presented in the CO condition. The indication here is that Arabic speakers rely more on consonants, which present lexical

information that allows them to access and retrieve lexical items, than on vowels, which convey structural information and which assist in top-down processing.

In summary, investigating S-Ratio was a useful experimental technique to show how segmental contributions differ in English and Arabic. We showed that consonantal information and vocalic information contribute equally to speech intelligibility in English, but their contributions were intermediated by other factors such as sentence-level cues. The magnitude of vocalic vs. consonantal information affects speech intelligibility on a similar basis. On the other hand, Arabic consonantal information is sensitive to S-Ratio manipulation (even if the segment reduction occurs to vowels), and speech intelligibility in turn is sensitive to consonantal information.

5.2.3. Speech Processing and Speech Intelligibility in Arabic vs. English

We have now developed a nuanced and multifaceted picture about segmental contributions in both systems, and we have been assuming that bottom-up vs. top-down processing is pertinent to segmental contributions, but we still need to understand how segmental contributions, processing, and intelligibility are amalgamated into a single theme.

In the current study, evidence suggests that in English sentences, sentential context allowed vowels to exceed consonants in terms of their contributions to speech intelligibility. Likewise, vowels in Arabic boosted speech intelligibility at the sentence level compared to the word level, an observation that is taken as evidence for some top-down processing at the sentence level. Hence, although it was not the goal of the

current study to evaluate top-down vs. bottom-up processing in concatenative vs. non-concatenative languages, this came as a byproduct; we can therefore interpret the overall results in this light.

For Arabic speakers, consonantal information appears to be used primarily for bottom-up processing at both the word and sentence levels, while vocalic information is secondarily used for top-down processing, mostly at the sentence level. For English speakers, both consonantal and vocalic information are used for bottom-up processing at both the word and sentence levels. Additionally, vocalic information is used secondarily for top-down processing, to some extent at the word level and to a greater extent at the word level. Figure 18 below depicts this interpretation, using the size of squares to indicate the extent of influence for each segment type (consonant vs. vowel), processing type (bottom-up vs. top-down), and speech type (word vs. sentence). The arrows' directions also represent the two directions for processing: bottom-up vs. top-down. Note that when one type of processing is associated with one type of segment, it is meant to be the main processing type. Of course, vowels and consonants can provide some information that can be used by listeners to process speech in a bottom-up manner when a top-down manner is stated, and vice versa, but it is not the dominant type of processing. For instance, we cannot claim that vowels do not provide any information for bottom-up processing at the word level in Arabic, but it is not the major type of processing because listeners use the other type of segments (consonants here) to process speech as well.

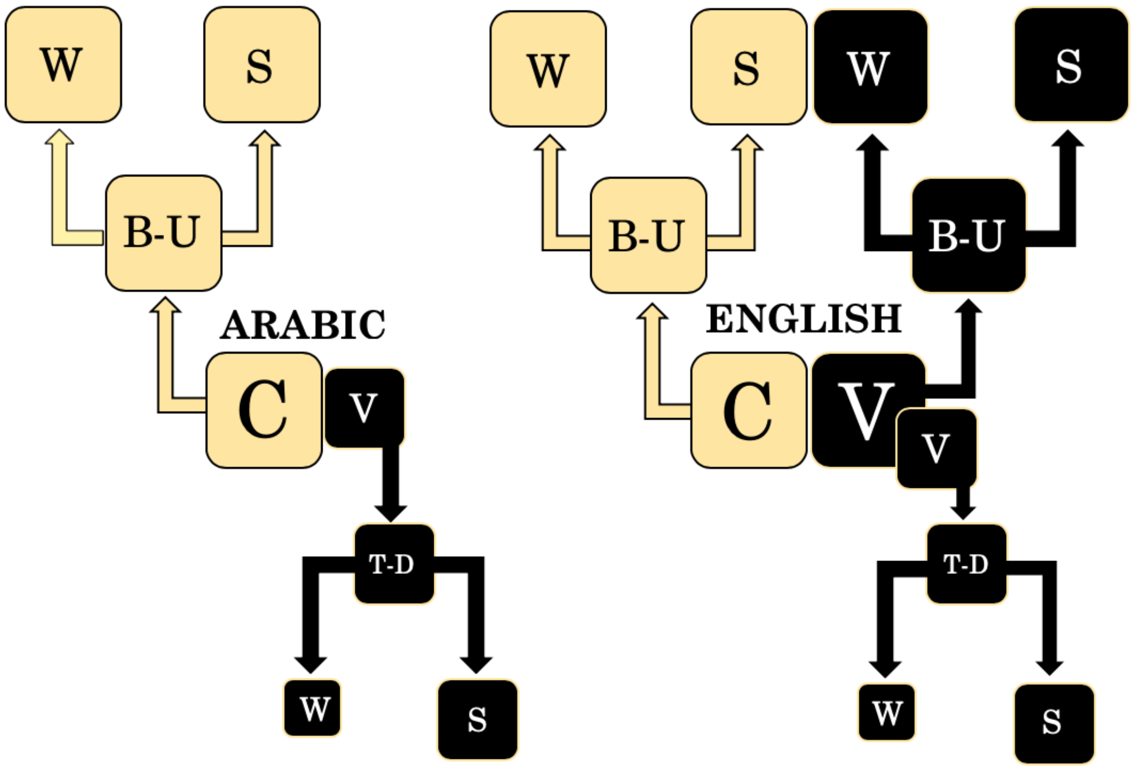


Figure 18. Illustration of speech processing in Arabic vs. English

Note: B-U: bottom-up, T-D: top-down, W: word, and S: sentence.

To summarize, the current results have shown that regardless of the diverse efforts in previous studies to determine whether speech is processed in a bottom-up or top-down manner, speech processing is essentially tied to segmental contributions. Speech intelligibility is ultimately based on both the type of segment presented to the listener and the type of processing mainly used for that specific segment, which in turn depends on the given language's morphological system.

5.3 Conclusions

From as early as 1929 to as recent as 2015, there has continued to exist a debate on segmental contributions with respect to which segment type contributes more to speech intelligibility within a specific language and cross-linguistically. Studies on languages such as English, Dutch, and Spanish exhibited similar results; overall there is a tendency for consonantal contributions to speech intelligibility to be greater at the word level and for vocalic contributions to be greater at the sentence level. The current study adds to the literature by investigating the issue in a language, namely Arabic, that has a distinctive morphological system predicted to mediate segmental contributions, and in English after reconsidering and resolving some critical design issues that had emerged in the literature. The study revealed novel findings about segmental contributions and S-Ratios not only in Arabic but also in English, which will have some theoretical and experimental implications on speech intelligibility.

5.3.1 Theoretical and Experimental Implications

The results from the current study bear some theoretical and experimental implications for research on speech intelligibility. First, although concatenative languages may not treat vowels and consonants as extremely different, they still make a distinction between vowels and consonants in structuring their morphemes. There are some instances, as in some Southern African languages (see e.g., Bleek, 1862), where vowels are used as prefixes or suffixes for case, gender, and (in)definite articles. If we survey a language (or languages) and find a substantial difference between the segments used for building roots (or stems) or for lexical meaning and those used

for affixation or morphosyntactic information, we may predict the findings from the current study to apply to such languages. It would be truly interesting if we find a language or a variety of language that prefers vowels for lexical information and can test if the opposite result will be borne out.

Second, theories such as the Cohort Model (Marslen-Wilson & Welsh, 1978) and Shortlist (Norris, 1994) treated vowels and consonants as interchangeable units. In these models, an initial segment activates some lexical candidates (a "cohort") and another eliminates some competitors. However, if vowels vs. consonants provide different amounts of information about the identity of a word, this raises a question as to whether vowels vs. consonants also play different roles in activating cohort candidates. If consonants are robust while vowels are prone to contextual effects, then we need to come up with a more detailed approach that incorporates such differential contributions. The effect of initial segment in our study *does* support the claim argued for in these models; the first segment presented to listeners is crucial to lexical activation, access, and recognition. However, as van Ooijen (1996) pointed out, these models just assumed that vowels and consonants are the same when it comes to candidate activation, as they did not distinguish between initial vowel and initial consonant. It would be a matter of interest to examine if initial vowels and consonants activate candidates and facilitate speech intelligibility equally, especially because a shortlist of words with initial vowels may not be as large as a shortlist of those with initial consonants in a particular language, or vice versa.

Third, it would be desirable to model segmental contributions to speech intelligibility universally, but the comparison between concatenative and nonconcatenative languages informs us that segmental contributions are language-specific. This was also supported by previous findings on Chinese; in both, vowels play a greater role in speech intelligibility in isolated words and full sentences. Therefore, future research should focus on factors that accumulate to determine which segment type contributes more to speech intelligibility. So far, spectral information, contextual information, morphological information, and perhaps tonal properties (in tone languages such as Chinese) have been shown to contribute to speech intelligibility. It would be of interest to explore more languages and find out if some other sources of information can be added to this list.

Fourth, speech intelligibility is based on the type of processing, which itself depends on the type of segment being processed. Segment type then determines its contribution, which is dependent on the morphological system the language utilizes as well as some additional factors such as initial segment and S-Ratio.

Fifth, the comparison here between English and Arabic, words and sentences, initial vowels and consonants, and balanced and imbalanced ratios yields some important observations about previous research. The current study therefore has some experimental implications, as summarized below.

- It has become clear that the use of words, especially CV and CVC items, leads to some misleading conclusions about segmental contributions, since initial

segment seems to play a significant role in speech intelligibility. Future research should consider this design aspect.

- Some previous researchers such as Kewley-Port (2007) assigned sentences for the VO condition that were different from those for the CO condition to different participants. The analysis of recordings for our study here has shown that some sentences, regardless of their identical design to the rest of the stimuli, were by far more intelligible to most participants than all other sentences, as indicated in Chapters 3 and 4. If their design was implemented in our study, those sentences would have biased the results in the condition to which they were assigned. Therefore, it is important for any future research to avoid this imperfect design.
- In the current study, I controlled over word frequency, word familiarity, and sentence naturalness to the extent possible. Studies of this type did not seem to emphasize this aspect. A future study that does not take into consideration this design facet may yield some unreliable results.
- On the basis of the fundamental differences between the balanced and imbalanced conditions in our study, it has become clear that S-Ratio is a key factor that a researcher must control for. Future studies must not ignore such a significant factor; otherwise, it may lead to artificial results.

5.3.2 Looking Backward and Looking Forward: Current Limitations and Future Directions

Although this dissertation made every attempt to control for variables that we thought could impact the results, there were still a couple of limitations that were unavoidable given the time available for this work to be completed.

- The initial segment in Arabic words caused some concerns about the conclusions drawn from the first experiment. Due to the structure of the Arabic morphological system, it was impossible to come up with truly root-pattern words that would start with vowels and serve as stimuli for the Arabic word-level experiments.
- For both Arabic and English sentences, the ideal design would have added a third condition in which vowels exceeded the number of consonants. Unfortunately, this was not possible, but the contrast between words and sentences, initial vowels and consonants, balanced and imbalanced ratios, and vowels-only and consonants-only may have compensated for the missing condition in order to understand speech intelligibility.
- Despite the fact that the number of participants in our study is higher than in many previous studies, a group of thirty participants would be ideal for a repeated-measure test because data points from one participant are by necessity likely to be correlated and hence function as a merely one data point.

As is often the case, the scope of the current study could not expand to investigate some relevant questions and therefore they were left for future research. Possible future directions are summarized below.

- The dominant morphological system in Arabic is nonconcatenative, but there is a portion of the Arabic lexicon that does not follow this morphological system. Such lexical items may be borrowed from other languages or are functional rather than content words. A follow-up study to examine this aspect of Arabic would be recommended.
- Since vowels are responsible for structural and suprasegmental information, it would be recommended to test differences in segmental contributions among the category of consonants. That is, consonants fall along a continuum, with stops being less similar to vowels than other segments such as glides are to vowels. Therefore, a plausible question is whether different sub-classes of consonants contribute to speech intelligibility equally.
- In the current design (as well as in previous studies), all vocalic and consonantal information was removed. It would be interesting to test speech intelligibility in partial CO vs. partial VO where some (but not all) consonants or vowels are replaced with silence.
- No errors were tolerated in the current scoring technique; therefore, even if the error involved only one segment, the entire word is deemed unintelligible. It would be of interest to re-score the responses in a more flexible way. For instance, we could apply the so-called "letter articulation" method (Fletcher,

1929) in which we score every word based on the recognition of each segment. A word like "cat" would be considered 33.33% intelligible if the subject reported "bat", for example.

- Related to this is the analysis of error patterns. One could further analyze the errors per word and find out what segments caused the rate intelligibility to be high or low.
- Finally, in our study we only had two categories of S-Ratios. It would be recommended to have multiple categories in which the number of segments increase segment by segment instead of implementing a 10-segment difference at once.

5.3.3 Recapitulation

This study investigated segmental contributions to speech intelligibility at the word level vs. the sentence level in Arabic vs. English representing two languages that have different morphological systems, nonconcatenative vs. concatenative, respectively. The outcome exhibited a difference in segmental contributions in the two languages. While consonants played a fundamentally greater role in speech intelligibility at both the word and sentence levels in Arabic, vowels and consonants mainly played equal roles in speech intelligibility at both levels in English.

⇐
Chapter
Five

REFERENCES

⇒
Appen-
dices

Abu-Rabia, S. (1997). Reading in Arabic orthography: The effect of vowels and context on reading accuracy of poor and skilled native Arabic readers. *Reading and Writing*, *9*(1), 65-78.

Allen, J. B. (1996). Harvey Fletcher's role in the creation of communication acoustics. *The Journal of the Acoustical Society of America*, *99*(4), 1825-1839.

Bastian, J., Delattre, P., & Liberman, A. M. (1959). Silent interval as a cue for the distinction between stops and semivowels in medial position. *The Journal of the Acoustical Society of America*, *31*(11), 1568-1568.

Benmamoun, E. (2003). The role of the imperfective template in Arabic morphology. *Language Acquisition and Language Disorders*, *28*, 99-114.

Bentin, S., & Feldman, L. B. (1990). The contribution of morphological and semantic relatedness to repetition priming at short and long lags: Evidence from Hebrew. *The Quarterly Journal of Experimental Psychology*, *42*(4), 693-711.

Bleek, W. H. I. (1862). *A comparative grammar of South African languages* (Vol. 1). London: Gregg International.

Bonatti, L. L., Pena, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations the role of consonants and vowels in continuous

speech processing. *Psychological Science*, *16*(6), 451-459.

Brame, M. K. (1970). *Arabic phonology: implications for phonological theory and historical Semitic* (Unpublished Doctoral dissertation). Massachusetts Institute of Technology, USA.

Burkle, T. Z. (2004). *Contribution of consonant versus vowel information to sentence intelligibility by normal and hearing-impaired listeners*. (Unpublished Master's thesis). Indiana University, USA.

Burkle, T. Z., Kewley-Port, D., Humes, L., & Lee, J. H. (2004). Contribution of consonant versus vowel information to sentence intelligibility by normal and hearing-impaired listeners. *Journal of Acoustic Society of America*. *115*, 2601.

Caramazza, A., & Hillis, A. E. (1991). Lexical organization of nouns and verbs in the brain. *Nature*, *349*(6312), 788.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, *28*(3), 297-332.

Chen, F., Wong, L. L., & Wong, E. Y. (2013). Assessing the perceptual contributions of vowels and consonants to Mandarin sentence intelligibility. *The Journal of the Acoustical Society of America*, *134*(2), EL178-EL184.

Chen, F., Wong, M. L. Y., Zhu, S., & Wong, L. L. N. (2015). Relative contributions of vowels and consonants in recognizing isolated Mandarin words. *Journal of Phonetics, 52*, 26-34.

Cole, R., Yan, Y., Mak, B., Fanty, M., & Bailey, T. (1996). The contribution of consonants versus vowels to word recognition in fluent speech: *International Conference on Acoustics, Speech, and Signal Processing (ICASSP'96)*. Atlanta, GA: IEEE.

Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M., & Gerstman, L. J. (1952). Some experiments on the perception of synthetic speech sounds. *The Journal of the Acoustical Society of America*, *24*(6), 597-606.

Cutler, A., Sebastián-Gallés, N., Soler-Vilageliu, O., & van Ooijen, B. (2000). Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. *Memory & Cognition*, *28*(5), 746-755.

Davies, M. (2017). *The corpus of contemporary American English (COCA): 560 million words, 1990-present*. Retrieved from https://corpus.byu.edu/coca/

Deutsch, A., & Frost, R. (2003). Lexical organization and lexical access in a non-concatenated morphology. *Language Acquisition and Language Disorder*, *28*, 165-186.

Deutsch, A., Frost, R., Pollatsek, A., & Rayner, K. (2000). Early morphological effects in word recognition in Hebrew: Evidence from parafoveal preview benefit. *Language and Cognitive Processes*, *15*(4-5), 487-506.

Diependaele, K., Grainger, J., & Sandra, D. (2012). Derivational morphology and skilled reading: An empirical overview. In M. J. Spivey, K. McRae & M. F.

Joanisse (Eds.), *The Cambridge handbook of psycholinguistics* (pp. 311-332). UK: Cambridge University Press.

Dudley, H. (1940). The carrier nature of speech. *Bell System Technical Journal*, *19*(4), 495-515.

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, *24*(6), 409-436.

Feldman, L. B. (1991). The contribution of morphology to word recognition. *Psychological Research*, *53*(1), 33-41.

Feldman, L. B. (2000). Are morphological effects distinguishable from the effects of shared meaning and shared form? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1431.

Feldman, L. B. (Ed.). (2013). *Morphological aspects of language processing*. UK: Psychology Press.

Feldman, L. B., & Moskovljević, J. (1987). Repetition priming is not purely episodic in origin. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(4), 573.

Feldman, L. B., Frost, R., & Pnini, T. (1995). Decomposing words into their constituent morphemes: Evidence from English and Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 947.

Fletcher, H. (1929). *Speech and Hearing*. New York: Van Nostrand.

Fogerty, D., & Humes, L. E. (2012). The role of vowel and consonant fundamental frequency, envelope, and temporal fine structure cues to the intelligibility of words and sentences. *The Journal of the Acoustical Society of America*, *131*(2), 1490-1501.

Fogerty, D., & Kewley-Port, D. (2009). Perceptual contributions of the consonant-vowel boundary to sentence intelligibility. *The Journal of the Acoustical Society of America*, *126*(2), 847-857.

Fogerty, D., Kewley-Port, D., & Humes, L. E. (2012). The relative importance of consonant and vowel segments to the recognition of words and sentences: Effects of age and hearing loss. *The Journal of the Acoustical Society of America*, *132*(3), 1667-1678.

Fu, Q. J., Zhu, M., & Wang, X. (2011). Development and validation of the Mandarin speech perception test. *The Journal of the Acoustical Society of America*, *129*(6), EL267-EL273.

Garofolo, J. S., Lamel, L. F., Fisher, W. M., Fiscus, J. G., & Pallett, D. S. (1993, February). *DARPA TIMIT acoustic-phonetic continuous speech corpus CD-ROM. NIST speech disc 1-1.1*. Retrieved from https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir4930.pdf

Gaskell, M. G., & Marslen–Wilson, W. D. (1999). Ambiguity, competition, and blending in spoken word recognition. *Cognitive Science*, *23*(4), 439-462.

Havy, M., Serres, J., & Nazzi, T. (2014). A consonant/vowel asymmetry in word-form processing: Evidence in childhood and in adulthood. *Language and Speech*, *57*(2), 254-281.

Holes, C. (2004). *Modern Arabic: Structures, functions, and varieties*. Washington, DC: Georgetown University Press.

Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye movement-contingent display change study. *Memory & Cognition*, *32*(4), 523-532.

Kewley-Port, D., Burkle, T. Z., & Lee, J. H. (2007). Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *122*(4), 2365-2375.

Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *The Journal of the Acoustical Society of America*, *29*(1), 98-104.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1-2), 71-102.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*(1),

29-63.

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 576.

Marslen-Wilson, W.D. (1984). Function and process in spoken word recognition: A tuto- rial review. In H. Bouma & D.G. Bouwhuis (Eds.), *Attention and performance X* (pp. 125–150). Hillsdale, NJ: Erlbaum.

Martinez, M. T., & Müllner, I. (2015, October). Specific exceptions driving variation: the role of orthography in modern Hebrew spirantization. *A Conference Paper Presented at the Annual Meetings on Phonology*, Vancouver, Canada.

Massaro, D. W. (1989). Testing between the TRACE model and the fuzzy logical model of speech perception. *Cognitive psychology*, *21*(3), 398-421.

Massaro, D. W., & Cohen, M. M. (1991). Integration versus interactive activation: The joint influence of stimulus and context in perception. *Cognitive Psychology*, *23*(4), 558-614.

Max, L., & Onghena, P. (1999). Some issues in the statistical analysis of completely randomized and repeated measures designs for speech, language, and hearing research. *Journal of Speech, Language, and Hearing Research*, *42*(2), 261-270.

McCall, R. B., & Appelbaum, M. I. (1973). Bias in the analysis of repeated-measures

designs: Some alternative approaches. *Child Development*, *44*(3), 401-415.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1-86.

McQueen, J. M., Norris, D., & Cutler, A. (1999). The time course of lexical involvement in phonetic categorization. *The Journal of the Acoustical Society of America*, *105*(2), 1398-1398.

Miller, J. L., & Eimas, P. D. (1995). Speech perception: From signal to word. *Annual Review of Psychology*, *46*(1), 467-492.

Mimouni, Z., Kehayia, E., & Jarema, G. (1998). The mental representation of singular and plural nouns in Algerian Arabic as revealed through auditory priming in agrammatic aphasic patients. *Brain and language*, *61*(1), 63-87.

Morton, J. (1969). Interaction of information in word recognition. *Psychological review*, *76*(2), 165.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech intelligibility. *Cognition*, *52*(3), 189-234.

Osborne, J. W. (2006). Bringing balance and technical accuracy to reporting odds ratios and the results of logistic regression analyses. *Practical Assessment, Research & Evaluation*, *11*(7), 1-6.

Owens, E., Talbott, C. B., & Schubert, E. D. (1968). Vowel discrimination of hearing-

impaired listeners. *Journal of Speech, Language, and Hearing Research*, *11*(3), 648-655.

Owren, M. J., & Cardillo, G. C. (2006). The relative roles of vowels and consonants in discriminating talker identity versus word meanings. *The Journal of the Acoustical Society of America*, *119*(3), 1727-1739.

Parkinson, D. B. (2006). *ArabiCorpus*. Retrieved from http://arabicorpus.byu.edu/

Ratcliffe, R. R. (1998). *The "broken" plural problem in Arabic and comparative Semitic: Allomorphy and analogy in non-concatenative morphology* (Vol. 168). Amsterdam: John Benjamins Publishing.

Ravid, D., & Shimron, J. (2003). A developmental perspective on root perception in Hebrew and Palestinian Arabic. *Language Acquisition and Language Disorders*, *28*, 293-320.

Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance* (Vol. 6). Hillsdale, NJ: Erlbaum.

Samuel, A. G. (1981). The role of bottom-up confirmation in the phonemic restoration illusion. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(5), 1124.

Sevald, C. A. (1996). *Evidence for the representation of syllables and syllable structure in the production of normal speech* (Unpublished Doctoral dissertation). University of Illinois at Urbana-Champaign, USA.

Sherbecoe, R. L., & Studebaker, G. A. (2004). Supplementary formulas and tables for calculating and interconverting speech recognition scores in transformed arcsine units. *International Journal of Audiology*, *43*(8), 442-448.

Shimron, J. (Ed.). (2003). *Language processing and acquisition in languages of Semitic, root-based, morphology* (Vol. 28). Amsterdam: John Benjamins Publishing.

Studebaker, G. A. (1985). A rationalized arcsine transform. *Journal of Speech, Language, and Hearing Research*, *28*(3), 455-462.

Studebaker, G. A., Bisset, J. D., Van Ort, D. M., & Hoffnung, S. (1982). Paired comparison judgments of relative intelligibility in noise. *The Journal of the Acoustical Society of America*, *72*(1), 80-92.

Studebaker, G.A., McDaniel, D.M., & Sherbecoe, R.L. (1995). Evaluating relative speech intelligibility performance using the proficiency factor and rationalized arcsine differences. *American Academy of Audiology*, *6*(2):173-82.

Szumilas, M. (2010). Explaining odds ratios. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *19*(3), 227.

Taft, M. (1985). The decoding of words in lexical access: A review of the morphographic approach. In D. Besner, T. G. Waller & G. E. MacKinnon (Eds.), *Reading research: Advances in theory and practice,* (Vol. 5). New York: Academic Press.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal* of *Verbal Learning and Verbal Behavior, 14,* 638447.

Toro, J. M., Nespor, M., Mehler, J., & Bonatti, L. L. (2008). Finding words and rules in a speech stream functional differences between vowels and consonants. *Psychological Science*, *19*(2), 137-144.

van Ooijen, B. (1996). Vowel mutability and lexical selection in English: Evidence from a word reconstruction task. *Memory & Cognition*, *24*(5), 573-583.

Watson, J. C. (2007). *The phonology and morphology of Arabic*. New York: Oxford University Press.

Wright, R., & Nichols, D. (2014). *Measuring Vowel Duration*. Retrieved from http://depts.washington.edu/phonlab/resources/measuring-duration.pdf

Zue, V. W., Seneff, S., & Glass, J. (1990). Speech database development at MIT: TIMIT and beyond. *Speech Communication*, *9*, 351–356.

⇐ Refer- ences

APPENDICES

APPENDIX A:

⇒ Curriculum Vitae

Stimuli in Experiment 1

	Word	Meaning	Type
1	/taabaʕ/	"Followed up"	Verb, 3C-3V
2	/ʃaarak/	"Participated"	Verb, 3C-3V
3	/laaħadˤ/	"Noticed"	Verb, 3C-3V
4	/zaman/	"Synchronized"	Verb, 3C-3V
5	/ħuduuθ/	"Occurrence"	Noun, 3C-3V
6	/wusˤuul/	"Arrival"	Noun, 3C-3V
7	/jaqiin/	"Certainty"	Noun, 3C-3V
8	/ʃabiih/	"Doppelganger"	Noun, 3C-3V
9	/fariħ/	"Cheered up"	Verb, 3C-2V
10	/tˤamiʕ/	"Greedily yearned"	Verb, 3C-2V
11	/salim/	"Survived"	Verb, 3C-2V
12	/ʔaðin/	"Permitted"	Verb, 3C-2V
13	/ʕumur/	"Age"	Noun, 3C-2V
14	/ħulum/	"Dream"	Noun, 3C-2V
15	/xuluq/	"Morals"	Noun, 3C-2V
16	/xumus/	"Fifth"	Noun, 3C-2V
17	/ʒaraa/	"Took place"	Verb, 2C-3V
18	/daʕaa/	"Called"	Verb, 2C-3V
19	/saʕaa/	"Endeavored"	Verb, 2C-3V
20	/sˤafaa/	"Became clear"	Verb, 2C-3V
21	/ridˤaa/	"Satisfaction"	Noun, 2C-3V
22	/ɣinaa/	"Wealth"	Noun, 2C-3V
23	/ribaa/	"Usury"	Noun, 2C-3V
24	/mina/	Mina (Tent City)	Noun, 2C-3V

APPENDIX B:

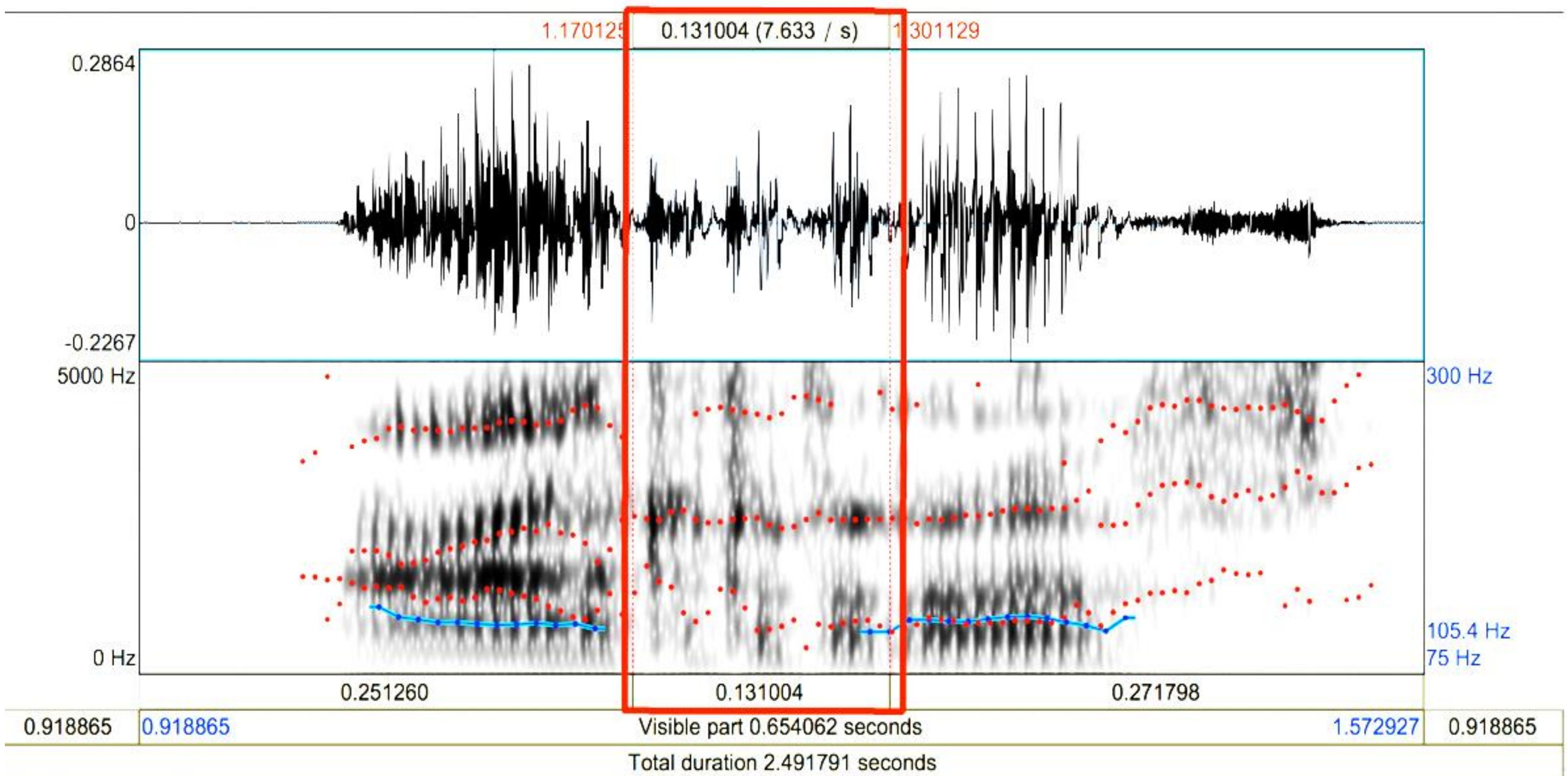
Word and Sentence Judgement Ratings in Experiments 1 and 2

	Written Form			Spoken Form		
Sentence	Sentence Naturalness	Sentence Likelihood	Word Familiarity	Sentence Naturalness	Sentence Likelihood	Word Familiarity
1*	6.61	6.73	6.69	6.43	6.95	6.95
2	6.61	6.58	7.00	6.44	7.00	7.00
3	6.61	6.87	7.00	6.69	6.82	7.00
4	6.32	6.40	6.95	6.69	6.82	7.00
5	6.50	7.00	6.95	6.73	7.00	7.00
6	6.80	6.94	6.94	6.71	7.00	7.00
7	6.60	7.00	6.96	6.79	6.75	6.75
8	6.66	7.00	7.00	6.79	6.89	6.89
9	6.47	7.00	7.00	6.79	6.89	6.89
10	6.99	7.00	7.00	6.81	7.00	7.00
11	6.85	6.86	7.00	6.81	7.00	7.00
12	6.85	6.86	6.83	6.77	6.82	6.82
13	6.50	7.00	6.95	6.51	6.82	6.82
14	6.45	7.00	6.94	6.79	7.00	7.00
15	6.61	7.00	6.87	6.81	7.00	7.00
16	6.99	7.00	7.00	6.83	7.00	7.00
17	6.80	6.94	7.00	6.91	6.95	6.95
18	6.85	7.00	6.86	6.91	6.95	6.95
19	6.99	7.00	7.00	6.91	6.95	6.95
20	6.85	6.86	6.83	6.91	6.95	6.95
21	6.80	6.94	6.94	6.92	7.00	7.00
22	6.79	7.00	6.99	6.51	6.87	7.00
23	6.99	7.00	7.00	6.61	6.90	6.90

24	6.85	6.86	6.83	6.76	7.00	7.00
25	6.99	7.00	7.00	6.69	6.80	7.00
26	6.99	7.00	7.00	6.75	6.70	7.00
27	6.99	7.00	7.00	6.51	6.87	6.87
28	6.61	6.73	7.00	6.46	6.85	6.85
29	6.99	7.00	7.00	6.83	6.92	6.92
30	6.80	6.94	6.94	7.00	7.00	7.00
31	6.80	6.94	6.94	6.81	6.81	6.81
32	6.66	6.78	7.00	7.00	7.00	7.00
33	6.70	7.00	6.98	6.80	6.96	6.96
34	6.56	7.00	6.96	7.00	7.00	7.00
35	6.70	7.00	6.98	7.00	7.00	7.00
36	6.80	7.00	6.99	7.00	6.80	7.00
37	6.70	6.83	7.00	7.00	6.85	7.00
38	6.50	6.90	6.89	7.00	7.00	7.00
39	6.50	7.00	6.95	6.71	6.93	7.00
40	6.85	6.86	7.00	6.57	6.89	6.89
41	6.70	7.00	6.98	6.71	6.93	7.00
42	6.66	6.78	6.75	6.81	6.96	6.96
43	6.59	7.00	6.96	6.51	6.87	7.00
44	6.80	7.00	7.00	6.71	6.93	7.00
45	6.66	6.78	6.75	6.60	6.90	6.90
46	6.32	6.40	7.00	6.81	7.80	7.00
47	6.60	7.00	6.96	7.00	6.90	6.90
48	6.45	7.00	6.94	7.00	6.85	7.00
Mean	6.70	6.91	6.95	6.77	6.93	6.96
	* Note that the sentences and words are ordered according the same order in Appendix A and I					

APPENDIX C:

Waveform and Spectrogram of the Arabic /r/



APPENDIX D:

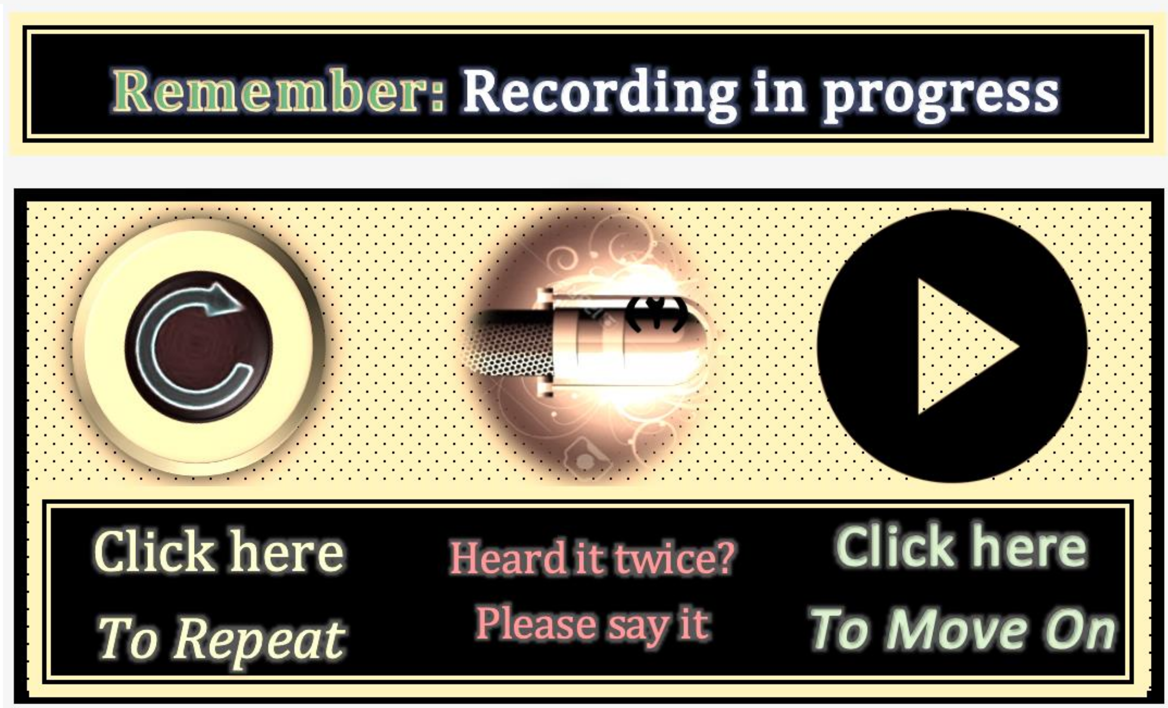
Segmental Durations in Arabic Stimuli in Experiments 1, 2, and 3

	C Durations	V Durations	C+V Durations
1*	0.136	0.130	0.133
2	0.148	0.154	0.151
3	0.134	0.139	0.137
4	0.133	0.139	0.136
5	0.144	0.179	0.161
6	0.140	0.157	0.149
7	0.146	0.161	0.153
8	0.147	0.153	0.150
9	0.163	0.165	0.164
10	0.137	0.141	0.139
11	0.127	0.144	0.135
12	0.146	0.169	0.158
13	0.153	0.164	0.158
14	0.154	0.156	0.155
15	0.146	0.169	0.158
16	0.138	0.173	0.155
17	0.145	0.172	0.159
18	0.121	0.164	0.142
19	0.133	0.154	0.144
20	0.129	0.173	0.151
21	0.140	0.154	0.147
22	0.129	0.173	0.151
23	0.126	0.154	0.140
24	0.140	0.181	0.160

25	0.139	0.138	0.139
26	0.166	0.156	0.161
27	0.153	0.149	0.151
28	0.188	0.158	0.173
29	0.160	0.164	0.162
30	0.156	0.148	0.152
31	0.143	0.164	0.153
32	0.171	0.147	0.159
33	0.139	0.155	0.147
34	0.155	0.159	0.157
35	0.173	0.163	0.168
36	0.147	0.180	0.163
37	0.161	0.141	0.151
38	0.162	0.170	0.166
39	0.148	0.181	0.165
40	0.164	0.182	0.173
41	0.147	0.181	0.164
42	0.155	0.166	0.161
43	0.134	0.160	0.147
44	0.131	0.153	0.142
45	0.140	0.179	0.160
46	0.130	0.167	0.149
47	0.154	0.198	0.176
48	0.132	0.166	0.149
Mean	0.146	0.161	0.152
	* Note that the sentences and words are ordered according the same order in Appendix A and I		

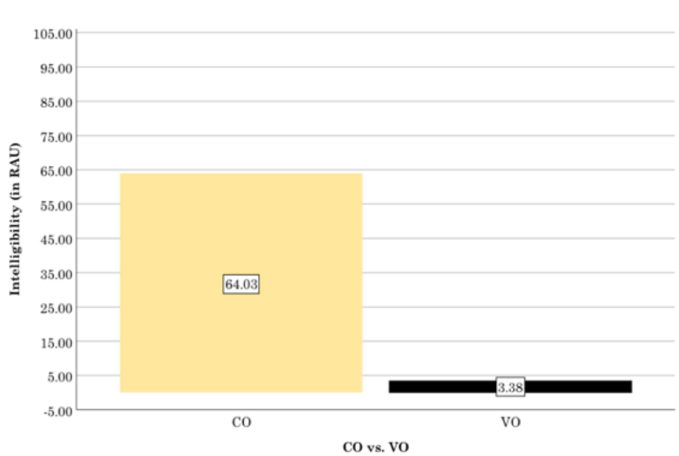
APPENDIX E:

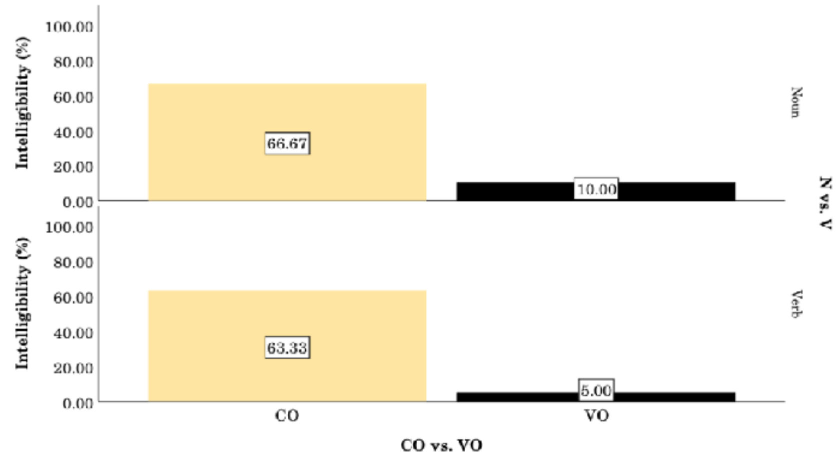
User Interfaces for Arabic and English Participants

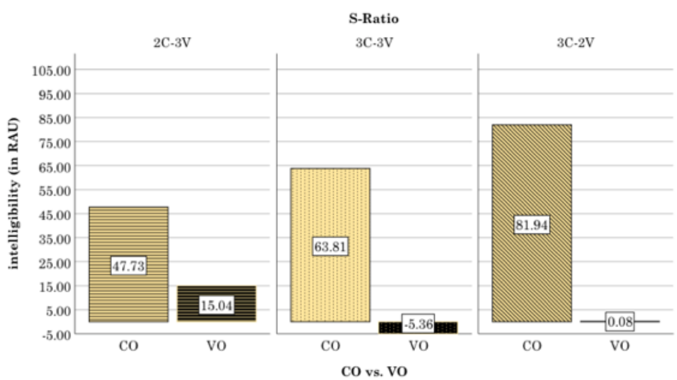


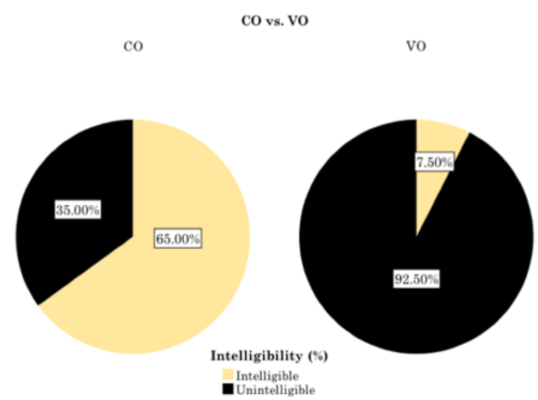
APPENDIX F:

Supplementary Display of Results in Experiment 1









APPENDIX G:

GEE Model Information in Experiment 1

Model Specifications			Selection	
Dependent Variable			Intelligibility	
Probability Distribution			Binomial	
Link Function			Logit	
Subject Effect & Repeated Effect			Subject & Trial	
Working Correlation Matrix Structure			Exchangeable	
Variables and Factors			N	Percent
Dependent Variable	Intelligibility	Unintelligible	153	63.7%
		Intelligible	87	36.3%
		Total	240	100.0%
Factor	CO vs. VO	CO	120	50.0%
		VO	120	50.0%
		Total	240	100.0%
	N vs. V	Noun	120	50.0%
		Verb	120	50.0%
		Total	240	100.0%
	S-Ratio	2C-3V	80	33.3%
		3C-3V	80	33.3%

	3C_2V	80	33.3%
	Total	240	100.0%

Data			Number
Levels	Subject Effect	Subject	20
	Repeated Effect	Trial	12
Subjects			20
Measurements per Subject	Minimum		12
	Maximum		12
Correlation Matrix Dimension			12

Log Quasi-Likelihood Function	Value [a]
Quasi Likelihood under Independence Model Criterion (QIC)[b]	223.56
Corrected Quasi Likelihood under Independence Model Criterion (QICC)[b]	221.37

Observed Category	Predicted Category		Correct Percentage
	Unintelligible	Intelligible	
Unintelligible	111	42	72.5%
Intelligible	9	78	89.7%
Overall			78.8%

APPENDIX H:

Complete GEE Output in Experiment 1

Parameter [b]	B	Std. Error	95% Wald CI		Hypothesis Test			Exp(B)	95% Wald (CI) for Exp(B)	
			Lower	Upper	Wald χ^2	*df*	Sig.		Lower	Upper
(Intercept)	-4.513	.7595	-6.002	-3.024	35.311	1	.000	.011	.002	.049
[CO-VO=1[c]]	4.878	.8421	3.227	6.528	33.546	1	**.000**	131.305	25.203	684.076
[CO-VO=0]	0[a]	.	.	.	.	.	.	1	.	.
[N-V=1]	1.326	.5859	.177	2.474	5.119	1	**.024**	3.765	1.194	11.872
[N-V=0]	0[a]	.	.	.	.	.	.	1	.	.
[S-Ratio=3C-32]	2.416	.9431	.568	4.264	6.563	1	**.010**	11.200	1.764	71.121
[S-Ratio=2C-3V]	1.062	1.4623	-1.804	3.928	.528	1	.468	2.892	.165	50.809
[S-Ratio=3C-3V]	0[a]	.	.	.	.	.	.	1	.	.
[CO-VO=1] x [N-V=1]	-.788	.5282	-1.823	.247	2.225	1	.136	.455	.161	1.281
[CO-VO=1] x [N-V=2]	0[a]	.	.	.	.	.	.	1	.	.
[CO-VO=0] x [N-V=1]	0[a]	.	.	.	.	.	.	1	.	.

[CO-VO=0] x [N-V=2]	0[a]	.	.	.	.	.	.	1	.	.
[CO-VO=1] x [S-Ratio=3C-2V]	-2.813	.9887	-4.751	-.875	8.092	1	**.004**	.060	.009	.417
[CO-VO=1] x [S-Ratio=2C-3V]	.091	1.5663	-2.979	3.161	.003	1	.954	1.095	.051	23.588
[CO-VO=1] x [S-Ratio=3C-3V]	0[a]	.	.	.	.	.	.	1	.	.
[CO-VO=0] x [S-Ratio=3C-2V]	0[a]	.	.	.	.	.	.	1	.	.
[CO-VO=0] x [S-Ratio=2C-3V]	0[a]	.	.	.	.	.	.	1	.	.
[CO-VO=0] x [S-Ratio=3C-3V]	0[a]	.	.	.	.	.	.	1	.	.
[N-V=0] x [S-Ratio=3C-2V]	-.667	.3204	-1.295	-.039	4.337	1	**.037**	.513	.274	.962
[N-V=1] x [S-Ratio=2C-3V]	-.462	.3137	-1.077	.153	2.170	1	.141	.630	.341	1.165
[N-V=1] x [S-Ratio=3C-3V]	0[a]	.	.	.	.	.	.	1	.	.
[N-V=0] x [S-Ratio=3C-2V]	0[a]	.	.	.	.	.	.	1	.	.
[N-V=0] x [S-Ratio=2C-3V]	0[a]	.	.	.	.	.	.	1	.	.
[N-V=0] x [S-Ratio=3C-3V]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1									

Dependent variable: Intelligibility Model: (Intercept), CO-VO, N-V, S-Ratio, CO-VO x N-V, CO-VO x S-Ratio, N-V x S-Ratio

a. Set to zero because this parameter is redundant. b. Reference category is: *Intelligible* c. CO is coded as 1, VO as 0, N as 1, and V as 0.

Parameter [b]	B	Std. Error	95% Wald Confidence Interval (CI)		Hypothesis Test			Exp(B)	95% Wald (CI) for Exp(B)	
			Lower	Upper	Wald χ^2	*df*	Sig.		Lower	Upper
(Intercept)	4.513	.7595	3.024	6.002	35.311	1	.000	91.193	20.583	404.035
[CO-VO=1[c]]	-4.878	.8421	-6.528	-3.227	33.546	1	**.000**	.008	.001	.040
[CO-VO=0]	0[a]	.	.	.	.	.	.	1	.	.
[N-V=1]	-1.326	.5859	-2.474	-.177	5.119	1	**.024**	.266	.084	.837
[N-V=0]	0[a]	.	.	.	.	.	.	1	.	.
[S-Ratio=3C-32]	-2.416	.9431	-4.264	-.568	6.563	1	**.010**	.089	.014	.567
[S-Ratio=2C-3V]	-1.062	1.4623	-3.928	1.804	.528	1	.468	.346	.020	6.073
[S-Ratio=3C-3V]	0[a]	.	.	.	.	.	.	1	.	.
[CO-VO=1] x [N-V=1]	.788	.5282	-.247	1.823	2.225	1	.136	2.199	.781	6.192
[CO-VO=1] x [N-V=2]	0[a]	.	.	.	.	.	.	1	.	.
[CO-VO=0] x [N-V=1]	0[a]	.	.	.	.	.	.	1	.	.
[CO-VO=0] x [N-V=2]	0[a]	.	.	.	.	.	.	1	.	.
[CO-VO=1] x [S-Ratio=3C-2V]	2.813	.9887	.875	4.751	8.092	1	**.004**	16.654	2.398	115.648
[CO-VO=1] x [S-Ratio=2C-3V]	-.091	1.5663	-3.161	2.979	.003	1	.954	.913	.042	19.666
[CO-VO=1] x [S-Ratio=3C-3V]	0[a]	.	.	.	.	.	.	1	.	.
[CO-VO=0] x [S-Ratio=3C-2V]	0[a]	.	.	.	.	.	.	1	.	.
[CO-VO=0] x [S-Ratio=2C-3V]	0[a]	.	.	.	.	.	.	1	.	.
[CO-VO=0] x [S-Ratio=3C-3V]	0[a]	.	.	.	.	.	.	1	.	.
[N-V=0] x [S-Ratio=3C-2V]	.667	.3204	.039	1.295	4.337	1	**.037**	1.949	1.040	3.651
[N-V=1] x [S-Ratio=2C-3V]	.462	.3137	-.153	1.077	2.170	1	.141	1.587	.858	2.936
[N-V=1] x [S-Ratio=3C-3V]	0[a]	.	.	.	.	.	.	1	.	.
[N-V=0] x [S-Ratio=3C-2V]	0[a]	.	.	.	.	.	.	1	.	.
[N-V=0] x [S-Ratio=2C-3V]	0[a]	.	.	.	.	.	.	1	.	.
[N-V=0] x [S-Ratio=3C-3V]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1									

Dependent variable: Intelligibility Model: (Intercept), CO-VO, N-V, S-Ratio, CO-VO x N-V, CO-VO x S-Ratio, N-V x S-Ratio

a. Set to zero because this parameter is redundant. b. Reference category is *Unintelligible* c. CO is coded as 1, VO as 0, N as 1, and V as 0.

APPENDIX I:

Sentences Used in Experiments 2 and 3

Sentences from Experiment 2

1-

taabaʕ[18]-a	fariiq-u	l-muħamaa-t-i	maa	qarrar-a-hu	l-muħaqqiq
follow up.PRF-3psm	team-NOM	the-law-F-GEN	what	determine.PRF-3psm.it	the-detective

"The law team followed up on what the detective determined/decided."

2-

ʃaarak-a	jaħjaa	ħamid-an	fii	muʔtamar-i	l-lisaanijjaat-i
participate.PRF.3psm	Yahya	Hamid-ACC	in	conference-GEN	the-linguistics-GEN

"Yahya participated with Hamid in the linguistics conference."

3-

laaħadˤ-a	badr-un	tˤaaʔir-at-an	ta-ʒuub-u	samaaʔ-a	l-madiin-ah
notice.PRF-3psm	Bader-NOM	plane-F-ACC	3psf-fly-IND	sky-ACC	the-city-F

"Bader noticed a plane flying in the city's sky."

4-

zaaman-a	tansˤiib-a	r-raʔiis-i	ʕadad-un	mina	l-mdˤahar-aat
coincide.PRF-3psm	inauguration-ACC	the-president-GEN	number-NOM	of	the- demonstration-P.F

18 Target words are in bold.

“A number of demonstrations coincided with the presidential inauguration.”

5-	qaatˤaʕ-a	l-ħudˤuur-u	ħadiith-a	l-masʔuul-i	laakinna-hu	**taabaʕ**
	interrupt.PRF-3psm	the-audience-NOM	speech-ACC	the-official-GEN	but-he	carry on.PRF

“The audience interrupted the official’s speech but he carried on.”

6-	lam	ja-ktariθ-i	l-fariiq-u	bi-l-xasaar-at-i	ħiinamaa	**ʃaarak**
	not.PRF	3psm-care-JUSS	the-team-NOM	about-the-loss-F-GEN	when	participate.PRF

“The team did not care about the [potential] loss when they participated.”

7-	inħaraf-at	sajjar-at-u	jaziid-a	fa-surʕaana	maa	**laaħaðˤ**
	swerve-3psf	car-F-NOM	Yaziid-GEN	then-quickly	as	notice.PRF

“Yaziid’s car swerved, but he (Yazid) quickly noticed [that].”

8-	xaalaf-a	li-ħtifaal-u	makaan-a	l-ʕardˤ-i	laakinna-hu	**zaaman**
	differ.PRF-3psm	the-celebration-NOM	place-ACC	the-show-GEN	but-it	coincide.PRF

“The celebration differed in location, but it [still] coincided with the show.”

9-	**ħuduuθ-u**	ħariiq-in	bi-l-mabnaa-Ø	ju-ʒib-u	xuruuʒ-a	s-sukkaan
	occurrence-NOM	fire-GEN	in-the-building-GEN	3psm-necessiate-IND	evacuation-ACC	the-residents

“The occurrence of fire in the building necessitates the residents' evacuation.”

10-**wusˤuul**-u	muħammad-in	madiin-at-a-naa	kaan-a	ʔaʒmal-a	xabar
arrival-NOM	Mohammed-GEN	city-F-ACC-our	be.PRF-3psm	best-ACC	news

“Mohammed’s arrival in our city was the best news.”

11-**jaqiin**-u	l-marʔ-i	bi-qudur-aat-i-hi	ja-ziid-u	min	saʔaad-at-i-h
certainty-NOM	the-individual-GEN	of-ability-P.F-GEN-his	3psm-increase-IND	of	happiness-F-GEN-his

“The individual’s certainty of his abilities increases his happiness.”

12-**ʃabiih**-u	kull-in	min-naa	tu-mayyiz-u-hu	saʒaajaa-Ø	mu-ʃtarak-ah
doppelganger-NOM	every-GEN	of-us	3psf-identify-IND-him	features-NOM	mutual-F

“For the doppelganger of every one of us, mutual features identify him.”

13-ja-ʒib-u	taʒhiiz-u	ʔadaw-aat-ii	li-t-taqdiim-i	qabl-a	l-**wusˤuul**
3psm-have to-IND	preparing-NOM	tool-P.F-my	for-the-presentation-GEN	before	the-arrival

“I have to prepare my tools for the presentation before the arrival.”

14-taħaddaθ-a	muħallil-uuna	ʕani	l-ħaadiθ-at-i	**duun-a**	**jaqiin**
talk.PRF-3psm	analyst-NOM.P.M	about	the-event-F-GEN	without-ACC	certainty

“Analysts talked about the event without being certain.”

15-ja-lzam-u	muʕaalaʒ-at-u	l-maʃaakil-i	ladaa	ʔawwal-i	**ħuduuθ**
3psm-must-IND	addressing-F-NOM	the-problems-GEN	at	first-GEN	occurrence

"Problems must be addressed at first occurrence."

16-hunaaka	maʕluum-aat-un	ʔuxraa	ladaa	masdˤar-in	**ʃabiih**
there	information-P.F-NOM	other	at	source-GEN	similar

"There is [some] other information from a similar source."

17-**fariħ**-a	rifaaq-u	ʕalijj-in	bi-naʒaaħ-i	mafʕuul-i	ʕilaaʒ-i-h
cheer.PRF-3psm	friends-NOM	Ali-GEN	for-success-GEN	effectiveness-GEN	cure-GEN-his

"Ali's friends cheered up for the success of his treatment effectiveness."

18-**tˤamiʕ**-a	ʔaʕdaaʔ-u	l-bilaad-i	fii	θaraw-aat-i-haa	l-ʕadiid-ah
greedily yearn.PRF-3psm	enemies-NOM	the-country-GEN	in	wealth-P.F-GEN-her	the-multiple-F

"The country's enemies [greedily] yearned for its multiple riches."

19-**salim**-a	qaaʔid-u	tˤ-tˤaaʔir-at-i	wa-tˤ-tˤaaqam-u	mina	l-ħaadiθ-ah
survive.PRF-3psm	pilot-NOM	the-plane-F-GEN	and-the-crew-NOM	from	the-incident-F

"The pilot and crew of the plane survived the incident."

20-**ʔaðin**-a	mudarrib-u	fariiq-i-naa	li-l-laaʕib-i	ʔan	ju-ɣaadir
permit.PRF-3psm	coach-NOM	team-GEN-our	for-the-player-GEN	to	3psm-leave

"Our team's coach permitted the player to leave."

21-fuuʒiʔ-a	kull-un	min-naa	bi-l-qaraar-aat-i	ʕaqibamaa	**fariħ**
surprise.PRF.Passive-3psm	everyone-NOM	of-us	by-the-decision-P.F-GEN	after	cheer up.PRF

"Everyone of us was surprised by the decisions after they cheered up."

22-larubbamaa	tafʃal-u	maʃaariʕ-u	l-ʔinsaan-i	ʔiðaa	**tˤamiʕ**
perhaps	3psf-fail-IND	projects-NOM	the-human-GEN	if	grow.PRF greedy

"A human's projects may fail if he/she grows greedy."

23-taʕarradˤ-a	qaarib-ii	li-huʒuum-in	ʔillaa	ʔanna-hu	**salim**
expose.PRF-3psm	boat-my	to-attack-GEN	but	that-it	survive.PRF

"My boat was exposed to an attack but it survived."

24-sa-ju-naffað-u	qaraar-u	l-qaadˤ-i	b-surʕ-at-in	ʔiðaa	ʔaðin
will-3psm-execute-IND	decision-NOM	the-judge-GEN	with-quickness-F-GEN	whenever	permit.PRF

"The judge's decision will be executed whenever he permits [that]."

25-**ʕumur**-u l-ʕatˤaaʔ-i laa ja-tawaqqaf-u ʕindamaa na-taaqaaʕad
age-NOM the-giving-GEN not 3psm-stop-IND when 1pp-retire
“The age of giving does not stop when we retire.”

26-**ħulum**-u l-ʕarab-i ħajaat-un biduun-i ħuruub-in wa-maʃaakil
dream-NOM the-Arabs-GEN life-NOM without-GEN wars-GEN and-problems
“The Arabs’ dream is a life without wars and problems.”

27-**xuluq**-u sˤ-sˤadiiq-i ju-qarrib-u wa-ju-baaʕid-u min-hu l-ʔaaxar-iin
moral-NOM the-friend-GEN 3psm-bring closer-IND and-3psm-distance-IND from-him the-other-ACC.P.M
“The friend’s morals bring closer or distance others from him.”

28-**xumus**-u tˤ-tˤullaab-i ʔasˤbaħ-a ju-ħibb-u ħudˤuur-a n-nadaw-aat
fifth-NOM the-students-GEN become.PRF-3psm 3psm-like-IND attending-ACC the-symposium-P.F
“A fifth of the students came to love attending symposiums.”

29-ʕaaʃ-a saalim-un ʃaɣaf-a l-lisaanijjaat-i tˤiilat-a l-**ʕumur**
live.PRF-3psm Salem-NOM passion-ACC the-linguistics-GEN throughout-ACC the-age/life
“Salem has had the passion for linguistics throughout his life.”

30-kullamaa	ja-muut-u	ʔamaam-a-kum	ʔamal-un	sa-ju-ulad-u	**ħulum**
every time	3psm-die-IND	in front-ACC-you	hope-NOM	will-3psm-bear.Passive-IND	dream

“Every time a hope dies, a dream gets born."

31-sˤadiiq-u-naa	ʕubd-u-llaah-i	kariim-un	wa-ʕalaa	ħusn-i	**xuluq**
friend-NOM-our	Abd-NOM-Allah-GEN	generous-NOM	and-upon	goodness-GEN	morals

“Our friend Abdullah is generous and with good morals.”

32-sa-ta-tadˤaaʔal-u	ʔarbaaħ-ii	min	θuluθ-in	ʔilaa	**xumus**
will-3psf-recede-IND	profits-my	from	third-GEN	to	fifth

“My profits will recede from a third to a fifth.”

33-**ʒaraa**	tadaawul-u	l-mustaʒadd-aat-i	maʕa	mabʕuuθ-i	l-ʔumam-i	l-muttaħid-ah
take place.PRF	discussing-NOM	the-update-P.F-GEN	with	envoy-GEN	the-nations-GEN	the-united-F

“Discussing the updates with the United Nations’ envoy took place.”

34-**daʕaa**	maʒlis-u	l-ʔamn-i	ʔilaa	mufaawadˤ-aat-in	mufaaʒʔ-ah
call.PRF	council-NOM	the-security-GEN	to	negotiation-P.F-GEN	emergent-F

“The Security Council called for urgent negotiations.”

35-**saʕaa** jaasir-un ʒaahid-an li-taqjiim-i ʔahdaaf-i l-kitaab
endeavor.PRF Yasser-NOM hard-ACC to-evaluating-GEN objectives-GEN the-book
“Yasser endeavored to evaluate the book’s objectives.”

36-**sˤafaa** ʒaww-u l-madiin-at-i fa-qadim-a ʔaalaaf-u l-musˤtˤaaf-iin
become clear.PRF weather-NOM the-city-F-GEN then-come.PRF -3psm thousands-NOM the-vacationer-GEN.P.M
“The weather of the city became clear and thousands of vacationers came.”

37-sa-ja-taʔakkad-u sˤ-sˤaħafiyy-uuna min ħaqiiq-at-i maa **ʒaraa**
will-3psm-ascertain-IND the-reporter-NOM.P.M of truth-F-GEN what occur.PRF
“The reporters will ascertain the truth of what occurred.”

38-kaan-a turkii mutatˤallʕ-an li-l-ħafl-at-i ħiinamaa **daʕaa**
be.PRF-3psm Turki looking forward-ACC to-the-party-F-GEN when invite.PRF
“Turki was looking forward to the party when he invited [us].”

39-sa-ju-ħaqqiq-u l-marʔ-u muʕðˤam-a tˤumuuħ-aat-i-hi ʔiðaa **saʕaa**
will-3psm-accomplish-IND the-one-NOM most-ACC ambition-P.F-GEN-his if endeavor.PRF
“One will accomplish his ambitions if he endeavors.”

40-tawallaa	l-muhandis-uuna	taʕqiim-a	l-maaʔ-i	ħattaa	**sˤafaa**
take care.PRF	the-engineer-NOM.P.M	sterilizing-ACC	the-water-GEN	until	become pure.PRF

“The engineers took care of sterilizing the water until it became pure.”

41-**ridˤaa**	malaajiin-i	l-mustaxdim-iina	ɣaajat-un	laa	tu-drak
satisfaction	millions-GEN	the-user-GEN.P.M	goal-NOM	not	3psf-reach.Passive

“The satisfaction of millions of users is a goal that is not reachable.”

42-**ɣinaa**	baʕdˤ-i	ʔunaas-in	ʒaʕal-a-hum	ja-ʕiib-uuna	l-fuqaraaʔ
wealth	some-GEN	people-GEN	make.PRF-3psm-them	3p-disgrace-pm	the-poor

“The wealth of some people made them disgrace the poor.”

43-**ribaa**	mullaak-i	l-bunuuk-i	ʔasˤbaħ-a	dˤaahir-at-an	xatˤiir-ah
usury	owners-GEN	the-banks-GEN	become.PRF-3psm	phenomenon-F-ACC	dangerous-F

“The usury of banks owners became a dangerous phenomenon.”

44-**minaa**	wa-muzdalifat-u	makaan-aani	fii	makkat-a	l-mukarram-ah
Mina	and-Muzdalifah-NOM	place-NOM.Dual	in	Mecca	the-venerable-F

“Mina and Muzdalifah are two places in the venerable Mecca.”

45-xuruuʒ-u	l-ʒumhuur-i	mubakkir-an	ʕalaamat-u	ʕadam-i	**ridˤaa**
leaving-NOM	the-audience-GEN	early-ACC	sign-NOM	lack-GEN	satisfaction

"The audience leaving [too] early is a sign of nonsatisfaction."

46-tˤabaq-aat-u	muʒtamaʕ-ii	ta-ʕiiʃ-u	faqr-an	wa-ta-ʕiiʃ-u	**ɣinaa**
stratum-F.P-NOM	society-my	3psf-live-IND	poverty-ACC	and-3psf-live-IND	wealth

"My society strata live in poverty and live in wealth."

47-tu-staʕaad-u	l-qurudˤ-u	l-ʕislamijj-at-u	duun-a	ʔajj-i	**ribaa**
3psf-repay.Passive-IND	the-loan-NOM	the-Islamic-F-NOM	without-ACC	any-GEN	usury

"Islamic loans are repaid without any usury."

48-ɣaadar-a	l-ħuʒʒaaʒ-u	maʃʕar-a	muzdalifat-a	tuʒaah-a	**minaa**
leave.PRF-3psm	the-pilgrims-NOM	Mashaaer-ACC	Muzdalifah-GEN	towards-ACC	Mina

"The pilgrims left Mashaaer Muzdalifah towards Mina."

Sentences from Experiment 3

1-	ʕaqd-u	n-nnaʃr-i	lam	ju-bdi-Ø	sm-a	muʔallif-i	ʃ-ʃiʕr
	contract-NOM	the-publication-GEN	not.PRF	3psm-show-JUSS	name-ACC	author-GEN	poem
	“The publication contract did not show the name of the poem author.”						

2-	ʕumr-u	l-marʔ-i	mulk-un	ja-ngusˤ-u	bi-fiʕl-i	l-waqt
	age-NOM	the-one-GEN	commodity-NOM	3psm-decrease-IND	with-function-GEN	the-time
	“One’s age is a commodity, which decreases as a function of time.”					

3-	lam	ja-kun	baʕdˤ-u	sˤ-sˤaħb-i	muktariθ-an	bi-l-ʔamr
	not.PRF	3psm-be	some-NOM	the-companions-GEN	concerned-ACC	with-the-matter
	“Some of the companions were not concerned about the matter.”					

4-	ʕind-a	l-matħaf-i	n-nusx-at-u	l-ʔaqdam-u	li-n-nasˤsˤ
	at-ACC	the-museum-GEN	the-copy-F-NOM	the-oldest-NOM	for-the-text
	“The museum has the oldest version of the text.”				

5-	ʔutruk-Ø	fi	l-manzil-i	ʃamʕ-at-an	ta-bʕaθ-u	d-ddifʔ
	leave-JUSS	in	the-house-GEN	candle-F-ACC	3psf-bring-IND	warmth
	“Leave, in the house, a candle to bring warmth.”					

6-	ʔinna	luʕb-at-a	ʃ-ʃatˤranʒ-i	qad	tu-nammi	l-ʕaql
	indeed	game-F-ACC	the-chess-GEN	may	3psf-develop	the-mind

“Indeed, chess may develop the mind.”

7-	lam	ja-kun-∅	mustaqbal-u	l-baħθ-i	li-ʕilm-i	l-luɣah
	not.PRF	3psm-be-JUSS	future-NOM	the-research-GEN	for-science-GEN	the-language

“The research future was not for language science.”

8-	ʔabhar-a	tˤ-tˤifl-u	ð-ðakijj-u	laʒn-at-a	l-ħukm
	impress.PRF.3psm	the-child-NOM	the-smart-NOM	committee-F-ACC	the-judgement

“The smart child impressed the judgement committee.”

9-	qad	ʔallaf-tu	qisˤsˤ-at-an	ʕan	riħl-at-i	l-ɣurb-ah
	already	write.PRF-1ps	story-F-ACC	about	journey-F-GEN	the-expatriation-F

“I have written a story about the expatriation journey.”

10-	qad	ta-ʒurr-u	l-ħarb-u	l-bald-at-a	li-wadˤʕ-in	sˤaʕb
	may	3psf-drag-IND	the-war-NOM	the-country-F-ACC	to-situation-GEN	difficult

“The war may drag the country into a difficult situation.”

11-lam	ta-ʃtari-Ø	l-ʔusr-at-u	manzil-an	ʕind-a	n-nahr
not.PRF	3psf-purchase-JUSS	the-family-F-NOM	house-ACC	at-ACC	the-river

“The family did not purchase a house by the river.”

12-ju-hdir-u	l-baʕdˤ-u	waqt-an	li-lbaħθ-i	ʕan	ʃaʔn-i-k
3psm-waste-IND	the-some-NOM	time-ACC	for-searching-GEN	about	affair-GEN-yours

“Some waste time searching about your affairs.”

13-ju-sˤbiħ-u	l-qamar-u	badr-an	ʕind-a	muntasˤaf-i	ʃ-ʃahr
3psm-beocme-IND	the-moon-NOM	full moon-ACC	at-ACC	middle-GEN	the-month

“The moon becomes a full moon in the middle of the month.”

14-baʕdˤ-u	l-ʔixwat-i	lam	ja-ʃmal-Ø-hum	daʕm-u	s-sakan
some-NOM	the-brothers-GEN	not.PRF	3psm-include-JUSS-them	subsidy-NOM	the-housing

“Some brothers, the housing subsidy did not include them.”

15-man	ja-ʕʃaq-Ø	taxasˤsˤusˤ-an	lam	ja-daʔ-Ø	taʔallum-a-h
who	3psm-love-JUSS	subject-ACC	not.PRF	3psm-stop-JUSS	learning-ACC-it

“Who loves a subject does not stop learning it.”

16-qad	ja-nqul-u	ʃ-ʃaʕb-u	sultˤ-at-a	l-ħukm-i	li-l-ħizb
may	3psm-transfer-IND	the-people-NOM	power-F-ACC	the-governance-GEN	to-the-party

"People may transfer the governance power to the party."

17-marr-at	tilka	l-biʕθ-at-u	qabl-a	ʔan	ʔastaslim
pass.PRF-3psf	that	the-scholarship-F-NOM	before-ACC	to	give up

"That scholarship [time] passed before I give up."

18-ʔaħdˤar-at	marjam-u	muntaʒ-an	li-faħsˤ-i	l-ʒism
bring.PRF-3psf	Maryam-NOM	product-ACC	for-inspection-GEN	the-body

"Maryam brought a product for body inspection."

19-ta-mlik-u	l-ʒadd-at-u	θarw-at-an	qad	ta-kfi	l-kull
3psf-own-IND	the-grandmother-f-NOM	wealth-F-ACC	may	3psf-suffice	the-everyone

"The grandmother owns a wealth that may suffice everyone."

20-lam	ja-ħdˤur-Ø	mohammid-un	ħafl-at-an	qabl-a	ʔams
not.PRF	3psm-attend-JUSS	Mohammed-NOM	party-F-ACC	before-ACC	yesterday

"Mohammed had not attended a party before yesterday."

21-lam	ta-bluɣ-Ø	maʒd-an	ʔumm-at-un	ta-ʔnaf-u	l-ʕamal
not-PRF	3psf-reach-JUSS	glory-ACC	nation-F-NOM	3psf-disdain-IND	the-work

"A nation that disdains work will not reach glory."

22-kun-Ø	qudw-at-an	ʔinna	n-nnaʃʔ-a	ja-tbaʕ-u	l-ʔakbar
be-JUSS	role model-F-ACC	indeed	the-youth-ACC	3psm-follow-IND	the-old

"Be a role model, [as] the youths follow the old."

23-tilka	l-ʔardˤ-u	tu-ʃbih-u	mustanqaʕ-an	li-l-qatl
that	the-land-NOM	resembles-IND	quagmire-ACC	for-the-murder

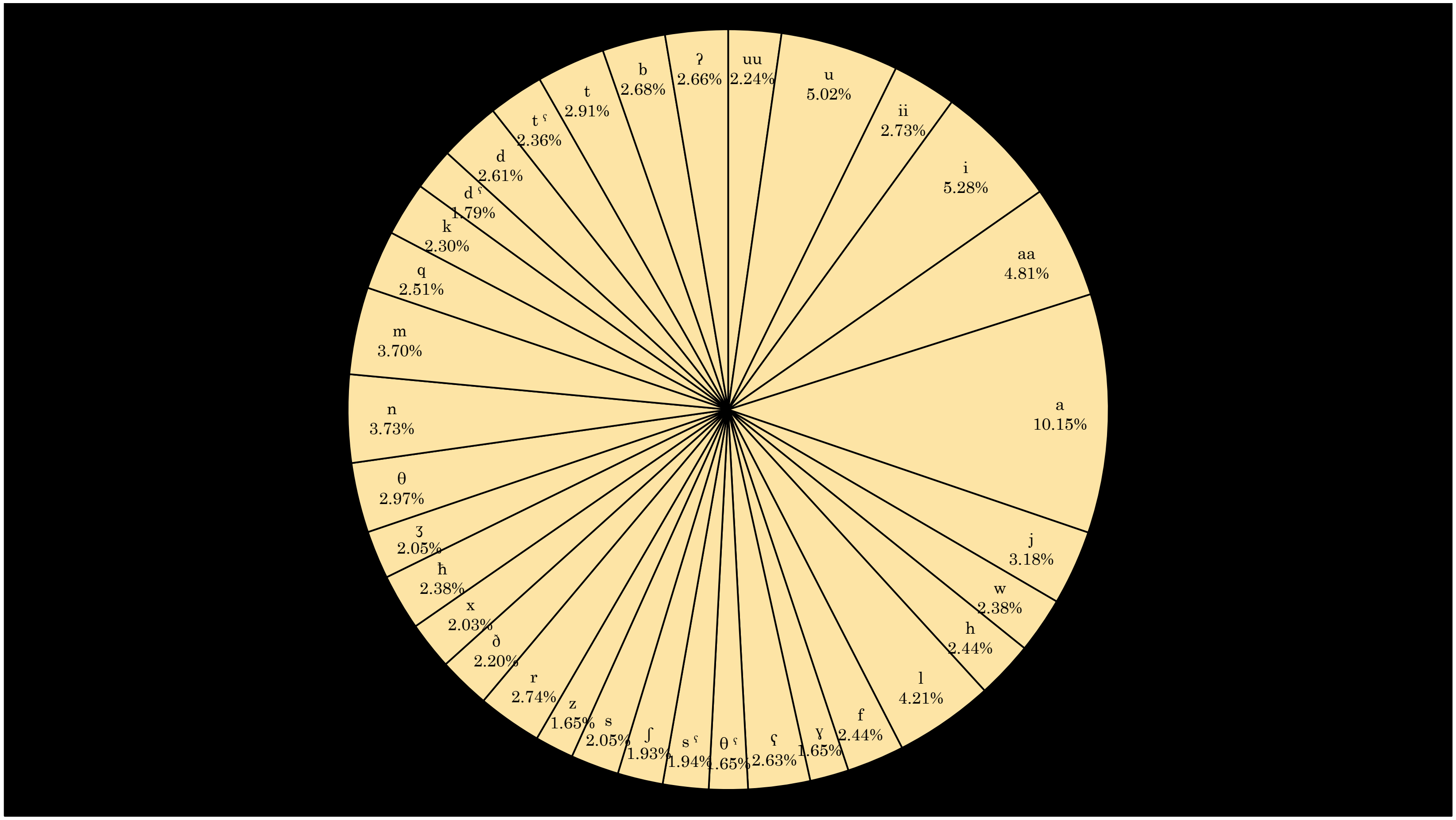
"That land resembles a quagmire of murder."

24-lamma	staʔad-tu	maʕ-a-ki	ðikra	l-ʔams-i	ʃtaq-t
when	recall.PRF-1ps	with-ACC-you	memory	the-past-GEN	yearn-1s

"When I recalled, with you, the past memories, I yearned [for you]"

APPENDIX J:

Proportions of Segments in the Arabic Stimuli



The pie chart shows that Arabic segments were almost equally frequent in the stimuli.
The /a/ vowel was more frequent than all other segments.

APPENDIX K:

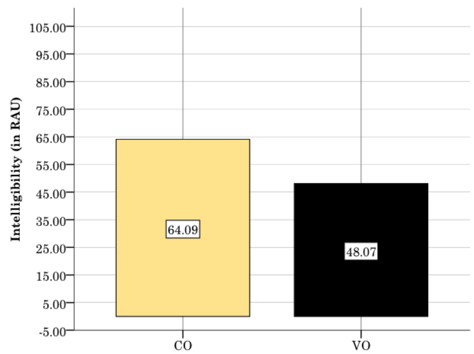
Segmental Duration and Speech Rate in Experiment 2

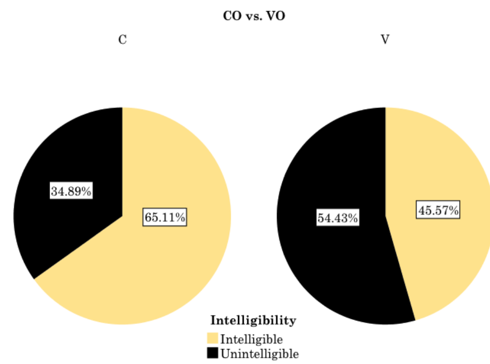
	C-Total Duration	V-Total Duration	C+ V-Total Duration	Phone per Second	Word per Second
1	2.72	2.47	5.19	8.87	0.86
2	2.95	2.62	5.57	8.26	0.93
3	3.08	2.51	5.59	8.23	0.93
4	2.80	2.50	5.29	8.69	0.88
5	3.17	3.22	6.39	7.20	1.06
6	3.22	2.83	6.05	7.60	1.01
7	3.21	2.90	6.11	7.53	1.02
8	3.23	2.75	5.98	7.69	1.00
9	3.58	2.80	6.37	7.22	1.06
10	3.01	2.68	5.69	8.09	0.95
11	2.91	2.74	5.65	8.14	0.94
12	3.21	3.22	6.43	7.16	1.07
13	3.51	2.95	6.47	7.11	1.08
14	3.24	2.97	6.21	7.40	1.04
15	3.20	3.22	6.42	7.16	1.07
16	3.17	2.94	6.11	7.52	1.02
17	3.18	3.28	6.46	7.12	1.08
18	2.77	2.95	5.72	8.04	0.95
19	2.79	2.94	5.73	8.03	0.95
20	2.70	3.29	5.99	8.01	1.00
21	2.94	2.78	5.71	8.05	0.95
22	2.96	2.94	5.90	7.80	0.98
23	2.53	2.77	5.30	8.68	0.88
24	3.07	3.25	6.33	7.27	1.05

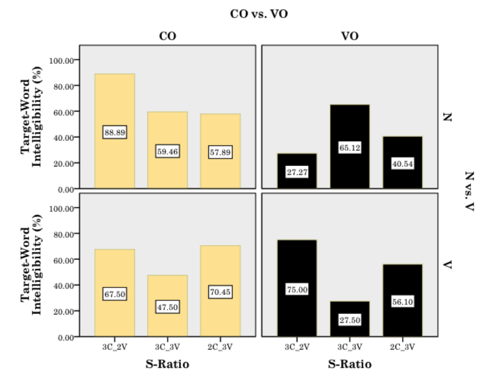
25	3.36	2.62	5.98	7.69	1.00
26	3.81	3.11	6.92	6.64	1.15
27	3.52	2.99	6.51	7.07	1.09
28	3.75	2.85	6.61	6.96	1.10
29	3.36	2.96	6.32	7.28	1.05
30	3.44	2.82	6.26	7.35	1.04
31	3.14	2.95	6.09	7.55	1.02
32	3.58	2.17	5.75	8.00	0.96
33	3.21	2.79	6.00	7.67	1.00
34	3.57	2.86	6.44	7.15	1.07
35	3.98	2.77	6.75	6.81	1.12
36	3.22	3.25	6.47	7.11	1.08
37	3.21	2.67	5.89	7.82	0.98
38	3.57	3.06	6.63	6.94	1.10
39	3.27	3.44	6.71	6.86	1.12
40	3.45	3.10	6.54	7.03	1.09
41	3.38	3.08	6.46	7.12	1.08
42	3.58	2.99	6.57	7.01	1.09
43	2.94	2.87	5.82	7.91	0.97
44	2.75	2.90	5.66	8.13	0.94
45	3.09	3.41	6.50	7.08	1.08
46	2.86	3.17	6.03	7.62	1.01
47	3.39	3.37	6.76	6.80	1.13
48	2.91	3.15	6.07	7.58	1.01
Mean	3.2	2.93	6.13	7.54	1.02
	* Note that the sentences and words are ordered according the same order in Appendix A and I				

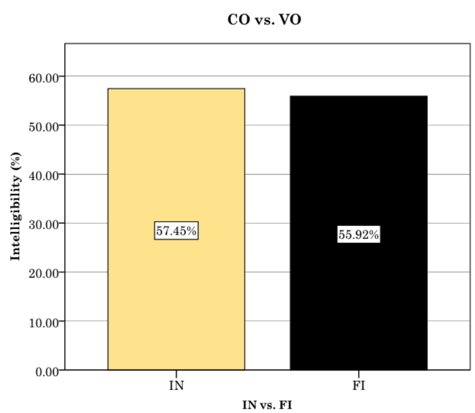
APPENDIX L:

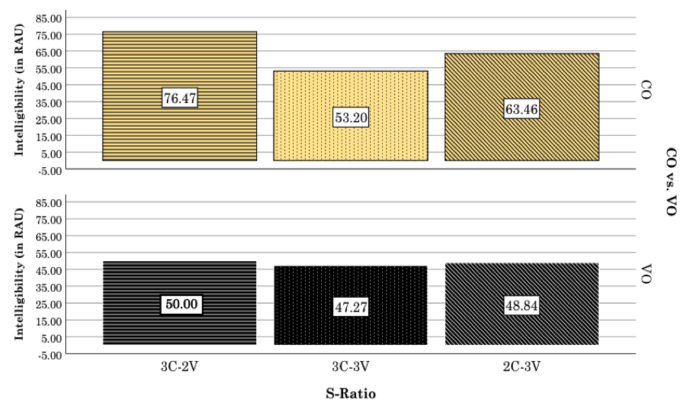
Supplementary Display of Results in Experiment 2 (Target-Word)

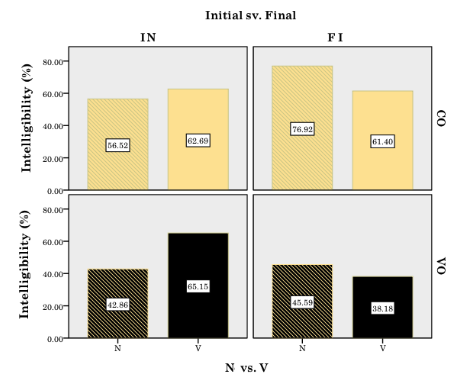












APPENDIX M:

GEE Model Information in Experiment 2 (Target-Word)

Model information			Selection
Dependent Variable			Intelligibility
Probability Distribution			Binomial
Link Function			Logit
Subject Effect & Within-Subject/Repeated Effect			Subject & Trial
Working Correlation Matrix Structure			Exchangeable
Data			**N**
Levels	Subject Effect	Subject	20
	Repeated Effect	Trial	24
Subjects			20
Measurements per Subject	Minimum		24
	Maximum		24
Correlation Matrix Dimension			24

Variables and Factors			N	Percent
Dependent Variable	Intelligibility	Unintelligible	208	63.7%
		Intelligible	272	36.3%
		Total	480	100.0%
Factor	CO vs. VO	CO	240	50.0%
		VO	240	50.0%

		Total	480	100.0%
	IN vs. FI	Noun	240	50.0%
		Verb	240	50.0%
		Total	480	100.0%
	N vs. V	Noun	240	50.0%
		Verb	240	50.0%
		Total	480	100.0%
	S-Ratio	2V-3C	160	33.3%
		3V-2C	160	33.3%
		3C-3V	160	33.3%
		Total	480	100.0%

Log Quasi-Likelihood Function	Value
Quasi Likelihood under Independence Model Criterion (QIC)[b]	628.852
Corrected Quasi Likelihood under Independence Model Criterion (QICC)	625.748

Observed Category	Predicted Category		Correct Percentage
	Unintelligible	Intelligible	
Unintelligible	121	85	58.73%
Intelligible	119	155	43.43%
Overall			57.50%

APPENDIX N:

Complete GEE Output in Experiment 2 (Target-Word)

			95% Wald CI		Hypothesis Test				95% Wald CI for Exp(B)	
Parameter	B	Std. Error	Lower	Upper	Wald χ^2	df	Sig.	Exp(B)	Lower	Upper
(Intercept)	-.218	.4163	-1.033	.598	.273	1	.601	.804	.356	1.819
[CO vs. VO=1.00]	1.487	.5528	.403	2.570	7.231	1	**.007**	4.422	1.496	13.068
[CO vs. VO=2.00]	0[a]	.	.	.	.	.	.	1	.	.
[NI vs. FI=2.00]	-.681	.3874	-1.440	.079	3.086	1	.079	.506	.237	1.082
[NI vs. FI=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[N vs. Verb=2.00]	.343	.3889	-.420	1.105	.776	1	.378	1.409	.657	3.018
[N vs. V=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[C_V Ratio=2.00]	.548	.3146	-.068	1.165	3.035	1	.081	1.730	.934	3.205
[C_V Ratio=1.00]	-.365	.4516	-1.250	.520	.655	1	.418	.694	.286	1.681
[C_V Ratio=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=1.00] x [NI vs. FI=2.00]	-.914	.3278	-1.557	-.272	7.777	1	**.005**	.401	.211	.762
[CO vs. VO=1.00] x [NI vs. FI=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=2.00] x [NI vs. FI=2.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=2.00] x [NI vs. FI=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=1.00] x [N vs. V=2.00]	-.532	.2565	-1.035	-.029	4.299	1	**.038**	.587	.355	.971

[CO vs. VO=1.00] x [N vs. V=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=2.00] x [N vs. V=2.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=2.00] x [N vs. V=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=1.00] x [C_V Ratio=2.00]	-.482	.3512	-1.171	.206	1.886	1	.170	.617	.310	1.229
[CO vs. VO=1.00] x [C_V Ratio=1.00]	.500	.4054	-.295	1.294	1.519	1	.218	1.648	.745	3.649
[CO vs. VO=1.00] x [C_V Ratio=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=2.00] x [C_V Ratio=2.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=2.00] x [C_V Ratio=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=2.00] x [C_V Ratio=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[NI vs. FI=2.00] x [N vs. V=2.00]	1.161	.4008	.375	1.946	8.390	1	**.004**	3.193	1.456	7.003
[NI vs. FI=2.00] x [N vs. V=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[NI vs. FI=3.00] x [N vs. V=2.00]	0[a]	.	.	.	.	.	.	1	.	.
[NI vs. FI=3.00] x [N vs. V=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[NI vs. FI=2.00] x [C_V Ratio=2.00]	.648	.3518	-.041	1.338	3.396	1	.065	1.912	.960	3.811
[NI vs. FI=2.00] x [C_V Ratio=1.00]	1.007	.3752	.272	1.742	7.203	1	**.007**	2.737	1.312	5.710
[NI vs. FI=2.00] x [C_V Ratio=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[NI vs. FI=3.00] x [C_V Ratio=2.00]	0[a]	.	.	.	.	.	.	1	.	.
[NI vs. FI=3.00] x [C_V Ratio=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[NI vs. FI=3.00] x [C_V Ratio=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[N vs. V=2.00] x [C_V Ratio=2.00]	-1.732	.3379	-2.394	-1.070	26.279	1	**.000**	.177	.091	.343
[N vs. V=2.00] x [C_V Ratio=1.00]	-.074	.4336	-.924	.776	.029	1	.864	.929	.397	2.172

[N vs. V=2.00] x [C_V Ratio=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[N vs. V=1.00] x [C_V Ratio=2.00]	0[a]	.	.	.	.	.	.	1	.	.
[N vs. V=1.00] x [C_V Ratio=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[N vs. V=1.00] x [C_V Ratio=3.00]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1									

Dependent variable: Target Word Intelligibility

Model: (Intercept), CO vs. VO, NI vs. FI, N vs. V, C_V Ratio, CO vs. VO x NI vs. FI, CO vs. VO x N vs. V, CO vs. VO x C_V Ratio, NI vs. FI x N vs. V, NI vs. FI x C_V Ratio, N vs. V x C_V Ratio

a. Set to zero because this parameter is redundant.

Parameter	B	Std. Error	95% Wald CI		Hypothesis Test			Exp(B)	95% Wald CI for Exp(B)	
			Lower	Upper	Wald χ^2	*df*	Sig.		Lower	Upper
(Intercept)	.218	.4163	-.598	1.033	.273	1	.601	1.243	.550	2.811
[CO vs. VO=1.00]	-1.487	.5528	-2.570	-.403	7.231	1	**.007**	.226	.077	.668
[CO vs. VO=2.00]	0[a]	.	.	.	.	.	.	1	.	.
[IN vs. FI=2.00]	.681	.3874	-.079	1.440	3.086	1	.079	1.975	.924	4.220
[IN vs. FI=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[N vs. V=2.00]	-.343	.3889	-1.105	.420	.776	1	.378	.710	.331	1.521
[N vs. V=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[C_V Ratio=2.00]	-.548	.3146	-1.165	.068	3.035	1	.081	.578	.312	1.071
[C_V Ratio=1.00]	.365	.4516	-.520	1.250	.655	1	.418	1.441	.595	3.492

[C_V Ratio=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=1.00] x [IN vs. FI=2.00]	.914	.3278	.272	1.557	7.777	1	**.005**	2.495	1.312	4.744
[CO vs. VO=1.00] x [IN vs. FI=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=2.00] x [IN vs. FI=2.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=2.00] x [IN vs. FI=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=1.00] x [N vs. V=2.00]	.532	.2565	.029	1.035	4.299	1	**.038**	1.702	1.030	2.814
[CO vs. VO=1.00] x [N vs. V=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=2.00] x [N vs. V=2.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=2.00] x [N vs. V=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=1.00] x [C_V Ratio=2.00]	.482	.3512	-.206	1.171	1.886	1	.170	1.620	.814	3.224
[CO vs. VO=1.00] x [C_V Ratio=1.00]	-.500	.4054	-1.294	.295	1.519	1	.218	.607	.274	1.343
[CO vs. VO=1.00] x [C_V Ratio=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=2.00] x [C_V Ratio=2.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=2.00] x [C_V Ratio=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=2.00] x [C_V Ratio=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[IN vs. FI=2.00] x [N vs. V=2.00]	-1.161	.4008	-1.946	-.375	8.390	1	**.004**	.313	.143	.687
[IN vs. FI=2.00] x [N vs. V=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[IN vs. FI=3.00] x [N vs. V=2.00]	0[a]	.	.	.	.	.	.	1	.	.
[IN vs. FI=3.00] x [N vs. V=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[IN vs. FI=2.00] x [C_V Ratio=2.00]	-.648	.3518	-1.338	.041	3.396	1	.065	.523	.262	1.042
[IN vs. FI=2.00] x [C_V Ratio=1.00]	-1.007	.3752	-1.742	-.272	7.203	1	**.007**	.365	.175	.762

[IN vs. FI=2.00] x [C_V Ratio=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[IN vs. FI=3.00] x [C_V Ratio=2.00]	0[a]	.	.	.	.	.	.	1	.	.
[IN vs. FI=3.00] x [C_V Ratio=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[IN vs. FI=3.00] x [C_V Ratio=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[N vs. V=2.00] x [C_V Ratio=2.00]	1.732	.3379	1.070	2.394	26.279	1	**.000**	5.652	2.915	10.959
[N vs. V=2.00] x [C_V Ratio=1.00]	.074	.4336	-.776	.924	.029	1	.864	1.077	.460	2.519
[N vs. V=2.00] x [C_V Ratio=3.00]	0[a]	.	.	.	.	.	.	1	.	.
[N vs. V=1.00] x [C_V Ratio=2.00]	0[a]	.	.	.	.	.	.	1	.	.
[N vs. V=1.00] x [C_V Ratio=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[N vs. V=1.00] x [C_V Ratio=3.00]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1									

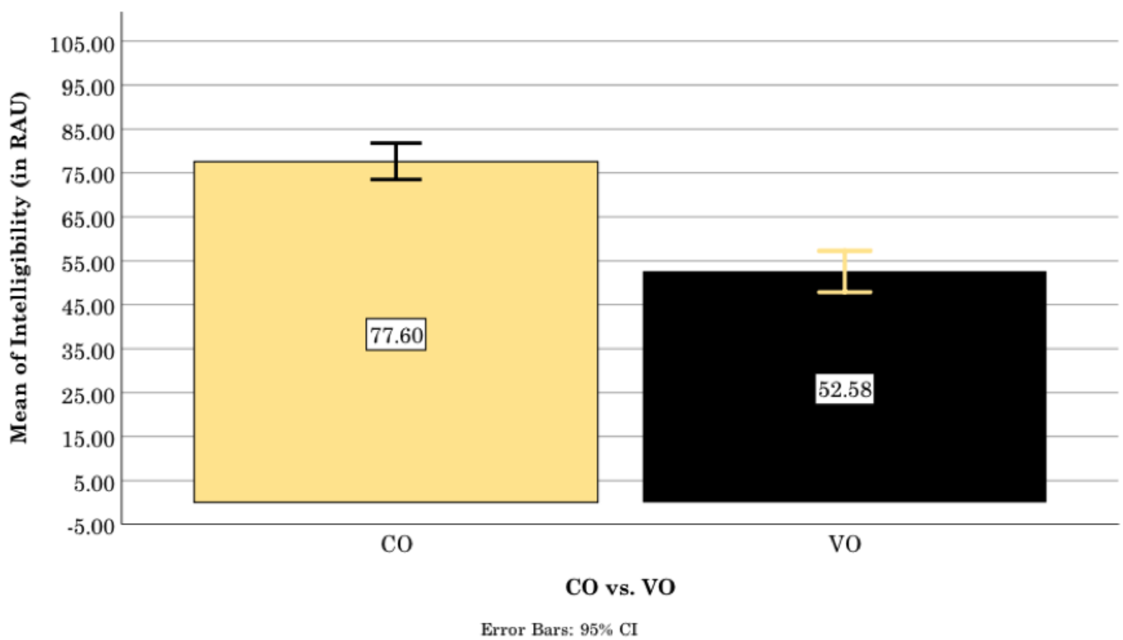
Dependent variable: Target Word Intelligibility

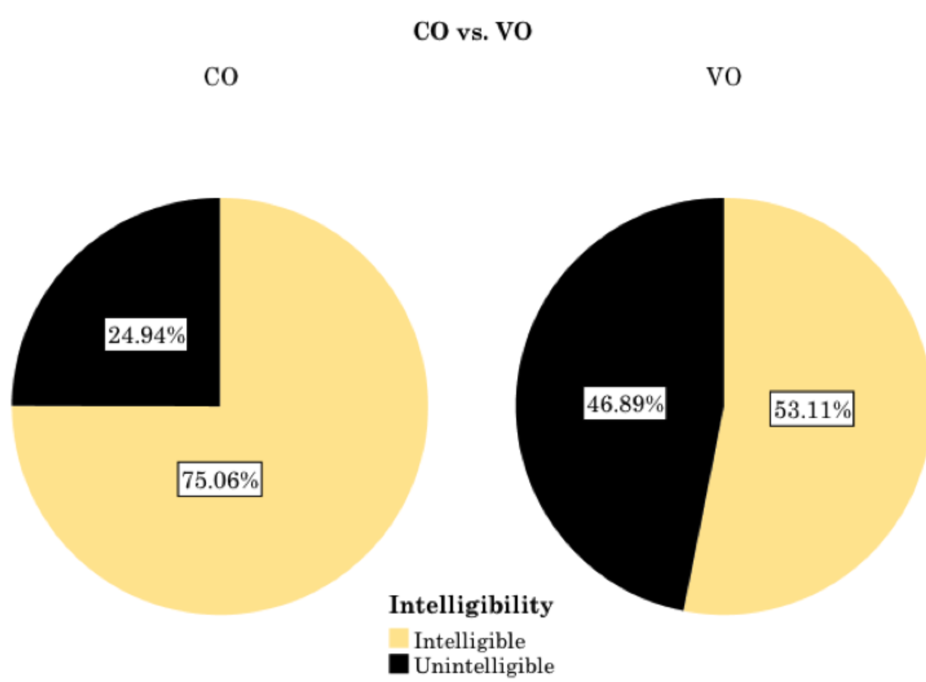
Model: (Intercept), CO vs. VO, IN vs. FI, N vs. V, S-Ratio, CO vs. VO x IN vs. FI, CO vs. VO x N vs. V, CO vs. VO x S-Ratio, IN vs. FI x N vs. V, IN vs. FI x S-Ratio, N vs. V x S-Ratio

a. Set to zero because this parameter is redundant.

APPENDIX O:

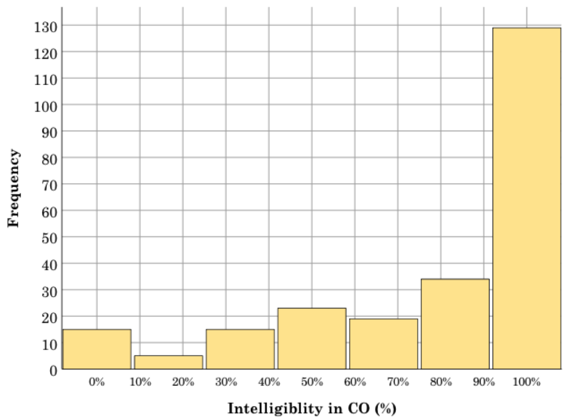
Supplementary Display of Results in Experiment 2 (Entire-Sentence)

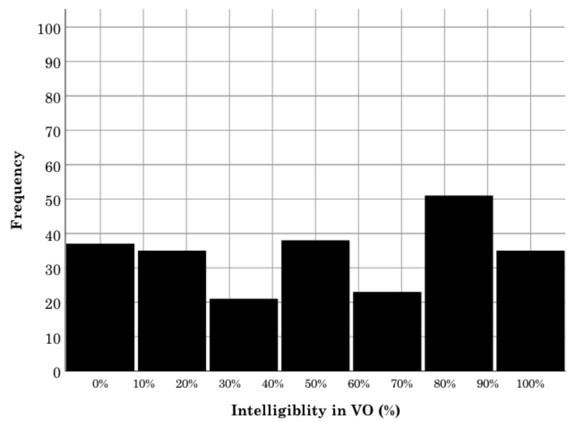


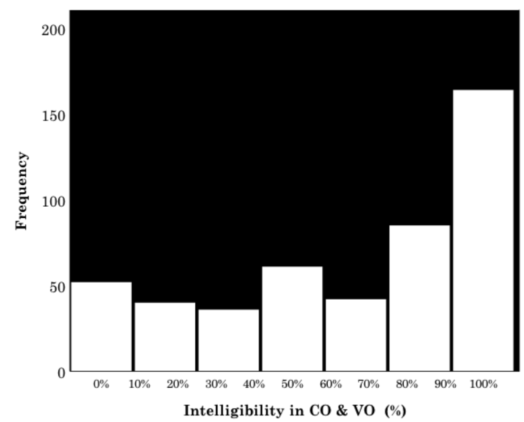


APPENDIX P:

Supplementary Display of Results in Experiment 2 (Frequencies)





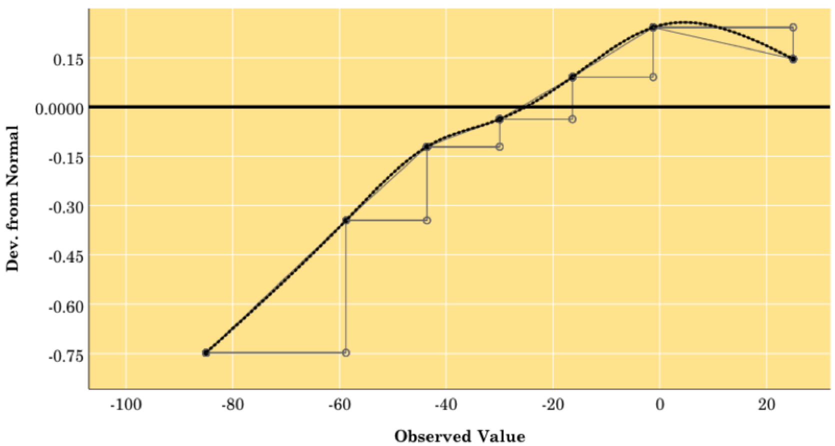


APPENDIX Q:

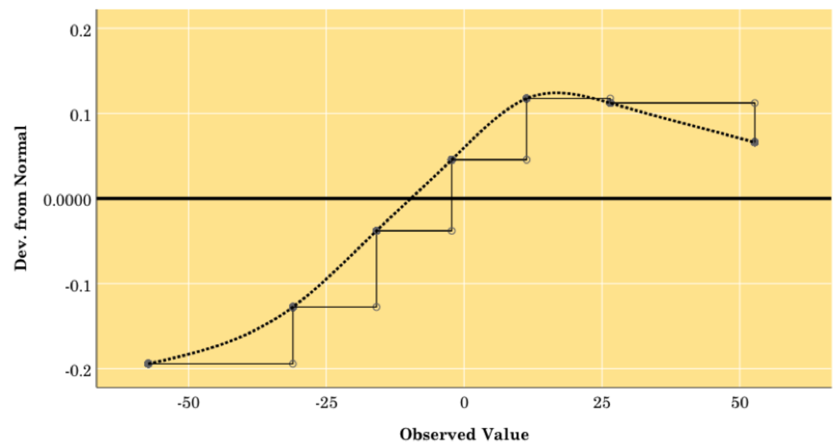
Normality Tests and Illustrations in Experiment 2 (Entire-Sentence)

CO vs. VO		Kolmogorov-Smirnov[a] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
RAU	CO	.314	240	**.000**	.760	240	**.000**
	VO	.144	240	**.000**	.921	240	**.000**

Condition / Value	Skewedness	Kurtosis
CO	-1.194	0.0491
VO	-0.155	-1.03



Detrended Q-Q Plot of Residuals for RAU in CO



Detrended Q-Q Plot of Residuals for RAU in VO

APPENDIX R:

GEE Model Information in Experiment 2 (Entire-Sentence)

Model Information		Selection
Dependent Variable		Intelligibility
Probability Distribution		Negative Binomial (MLE)
Link Function		Log
Subject Effect & Within-Subject/Repeated Effect		Subject & Trial
Working Correlation Matrix Structure		Exchangeable
Data		**Number**
Levels	Subject Effect Subject	20
	Repeated Effect Trial	24
Subjects		20
Measurements per Subject	Minimum	24
	Maximum	24
Correlation Matrix Dimension		24

Factor		Number	Percent
CO vs. VO	CO	240	50.0%
	VO	240	50.0%
	Total	480	100.0%

Dependent Variable	N	Minimum	Maximum	Mean	Std.
Intelligibility (Words per Sentence)	480	0	6	3.90	2.110

Log Quasi-Likelihood Function	Value
Quasi Likelihood Under Independence Model Criterion (QIC)[b]	703.282
Corrected Quasi Likelihood Under Independence Model Criterion (QICC)	693.387

APPENDIX S:

Mean Words per Sentence in Arabic and English Stimuli

Number	English		Arabic	
	Balanced	Imbalanced	Balanced	Imbalanced
1	6	5	6	7
2	7	5	6	6
3	6	6	6	6
4	5	6	6	5
5	6	6	6	6
6	6	5	6	6
7	6	5	6	6
8	7	5	6	5
9	6	5	6	6
10	7	5	6	6
11	6	6	6	6
12	6	5	6	6
13	7	6	6	6
14	6	6	6	6
15	6	6	6	6
16	6	6	6	6
17	7	5	6	6

18	5	6	6	5
19	5	5	6	6
20	6	6	6	6
21	6	5	6	6
22	7	5	6	6
23	5	5	6	5
24	6	5	6	6
Mean	**6.08**	**5.41**	**6**	**5.87**
STD	**0.65**	**0.50**	**0**	**0.44**
Mean	**5.70833333**		**5.87**	
STD	**0.55931445**		**0.22**	

* Note that the sentences and words are ordered according the same order in Appendix A and I

APPENDIX T:

Word and Sentence Judgement Ratings in Experiment 3

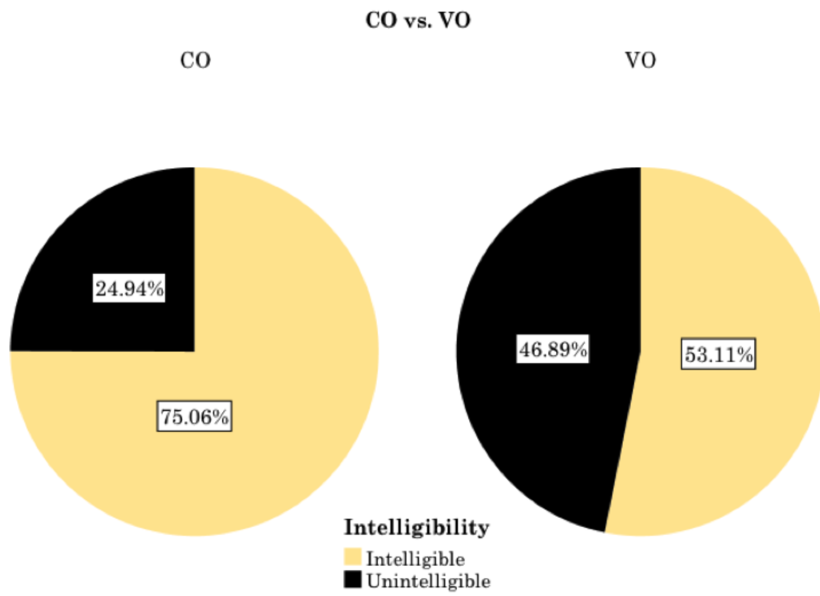
Sentence	Written Form	Spoken Form
1*	6.10	6.00
2	6.60	6.70
3	6.90	6.20
4	6.40	6.70
5	6.30	6.70
6	6.20	6.10
7	6.30	6.80
8	7.00	6.40
9	6.40	6.80
10	6.40	6.30
11	6.40	6.50
12	6.70	6.40
13	6.90	6.50
14	6.70	6.40
15	6.40	6.20

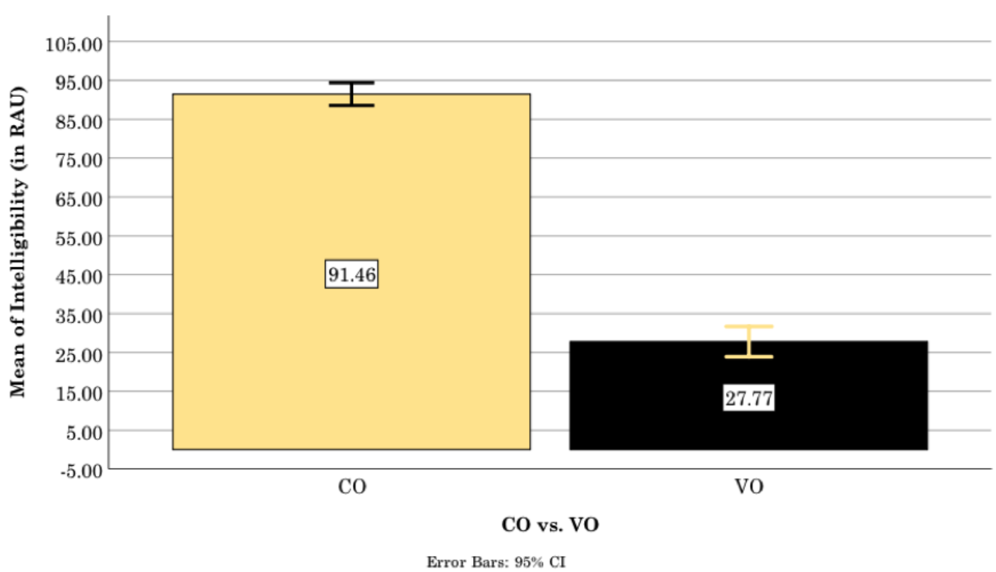
16	6.40	6.00
17	6.80	6.30
18	6.60	6.50
19	6.70	5.70
20	6.30	6.20
21	6.50	6.60
22	6.10	6.20
23	6.70	6.10
24	6.50	6.00
Mean	**6.51**	**6.35**
SD	**0.24901982**	**0.2888947**

* Note that the sentences and words are ordered according the same order in Appendix A and I

APPENDIX U:

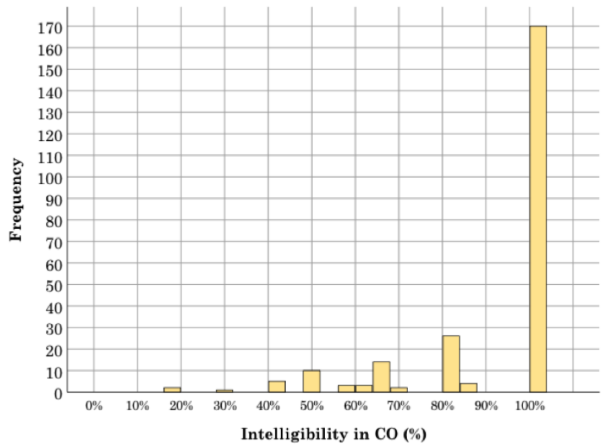
Supplementary Display of Results in Experiment 3

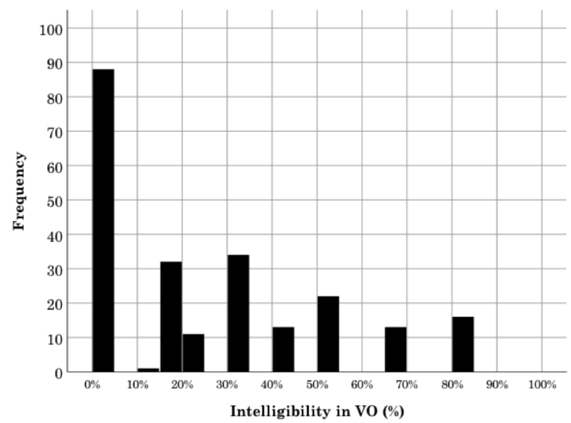


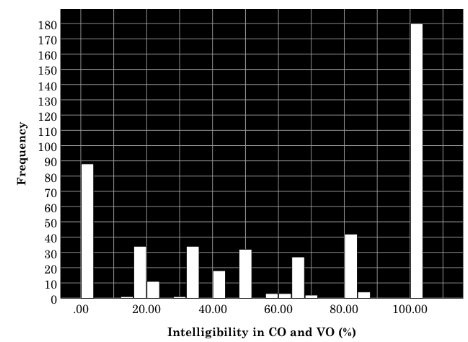


APPENDIX V:

Supplementary Display of Results in Experiment 3 (Frequencies)





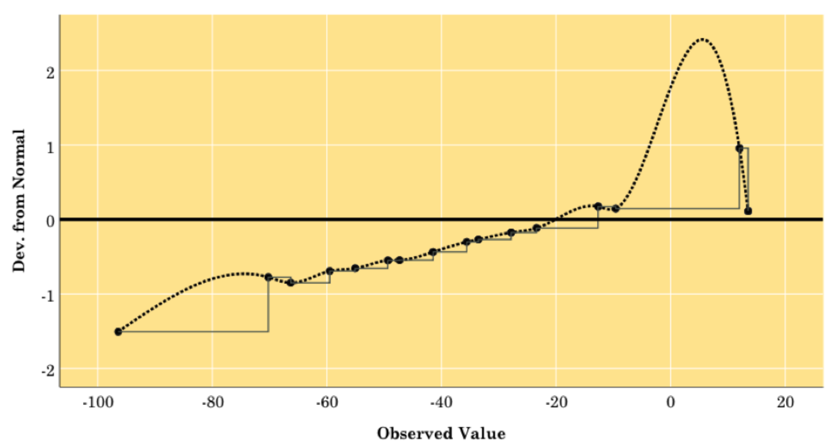


APPENDIX W:

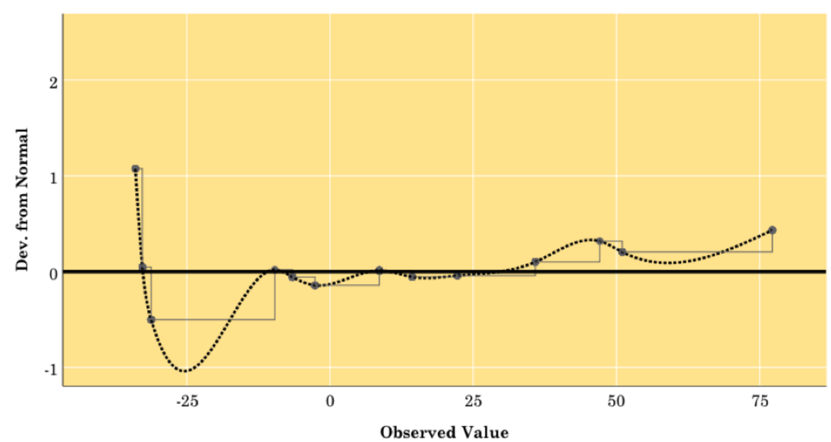
Normality Tests and Illustrations in Experiment 3

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
CO vs. VO		Statistic	*df*	Sig.	Statistic	*df*	Sig.
RAU	CO	.401	240	**.000**	.644	240	**.000**
	VO	.212	240	**.000**	.885	240	**.000**

Condition / Value	Skewedness	Kurtosis
CO	-1.647	-1.548
VO	0.607	-0.442



Detrended Q-Q Plot of Residuals for RAU in CO



Detrended Q-Q Plot of Residuals for RAU in VO

APPENDIX X:

GEE Model Information in Experiment 3

Model Information			Selection
Dependent Variable			Intelligibility
Probability Distribution			Negative Binomial (MLE)
Link Function			Log
Subject Effect & Within-Subject/Repeated Effect			Subject & Trial
Working Correlation Matrix Structure			Exchangeable
Data			**N**
Levels	Subject Effect	Subject	20
	Repeated Effect	Trial	24
Subjects			20
Measurements per Subject	Minimum		24
	Maximum		24
Correlation Matrix Dimension			24
Factor		**N**	**Percent**

CO vs. VO	CO	240	50.0%
	VO	240	50.0%
	Total	480	100.0%

Dependent Variable	N	Minimum	Maxi-mum	Mean	Std.
Intelligibility (number of words)	480	.00	6.00	3.4937	2.36713

Log Quasi-Likelihood Function	Value
Quasi Likelihood Under Independence Model Criterion (QIC)[b]	582.792
Corrected Quasi Likelihood Under Independence Model Criterion (QICC)	572.303

APPENDIX Y:

List of Words Used in Experiment 4

	Word	S-Ratio	Initial Segment	Experiment
1	Delicacy	4C-4V	C	Balanced
2	Vitality	4C-4V	C	Balanced
3	Missionary	4C-4V	C	Balanced
4	Monopoly	4C-4V	C	Balanced
5	Notify	3C-3V	C	Balanced
6	Jealousy	3C-3V	C	Balanced
7	Cavity	3C-3V	C	Balanced
8	Modify	3C-3V	C	Balanced
9	Goofy	2C-2V	C	Balanced
10	Shady	2C-2V	C	Balanced
11	Decay	2C-2V	C	Balanced
12	Woody	2C-2V	C	Balanced
13	Analytic	4C-4V	V	Balanced
14	Insanity	4C-4V	V	Balanced
15	Operative	4C-4V	V	Balanced
16	Intimacy	4C-4V	V	Balanced
17	Automate	3C-3V	V	Balanced
18	Elevate	3C-3V	V	Balanced
19	Allocate	3C-3V	V	Balanced

20	Infamy	3C-3V	V	Balanced
21	Awash	2C-2V	V	Balanced
22	Edit	2C-2V	V	Balanced
23	Akin	2C-2V	V	Balanced
24	Attach	2C-2V	V	Balanced
25	Thrift	4C-1V	C	Imbalanced
26	Strait	4C-1V	C	Imbalanced
27	Tract	4C-1V	C	Imbalanced
28	Drift	4C-1V	C	Imbalanced
29	Tempt	4C-1V	C	Imbalanced
30	Trench	4C-1V	C	Imbalanced
31	Shrink	4C-1V	C	Imbalanced
32	Scratch	4C-1V	C	Imbalanced
33	Bled	3C-1V	C	Imbalanced
34	Wink	3C-1V	C	Imbalanced
35	Bulb	3C-1V	C	Imbalanced
36	Spit	3C-1V	C	Imbalanced
37	Slug	3C-1V	C	Imbalanced
38	Slab	3C-1V	C	Imbalanced
39	Mint	3C-1V	C	Imbalanced
40	Lend	3C-1V	C	Imbalanced
41	Glue	2C-1V	C	Imbalanced

42	Hen	2C-1V	C	Imbalanced
43	Vet	2C-1V	C	Imbalanced
44	Mug	2C-1V	C	Imbalanced
45	Flu	2C-1V	C	Imbalanced
46	Numb	2C-1V	C	Imbalanced
47	Zip	2C-1V	C	Imbalanced
48	Shrew	2C-1V	C	Imbalanced

APPENDIX Z:

Judgement Ratings in Experiment 4

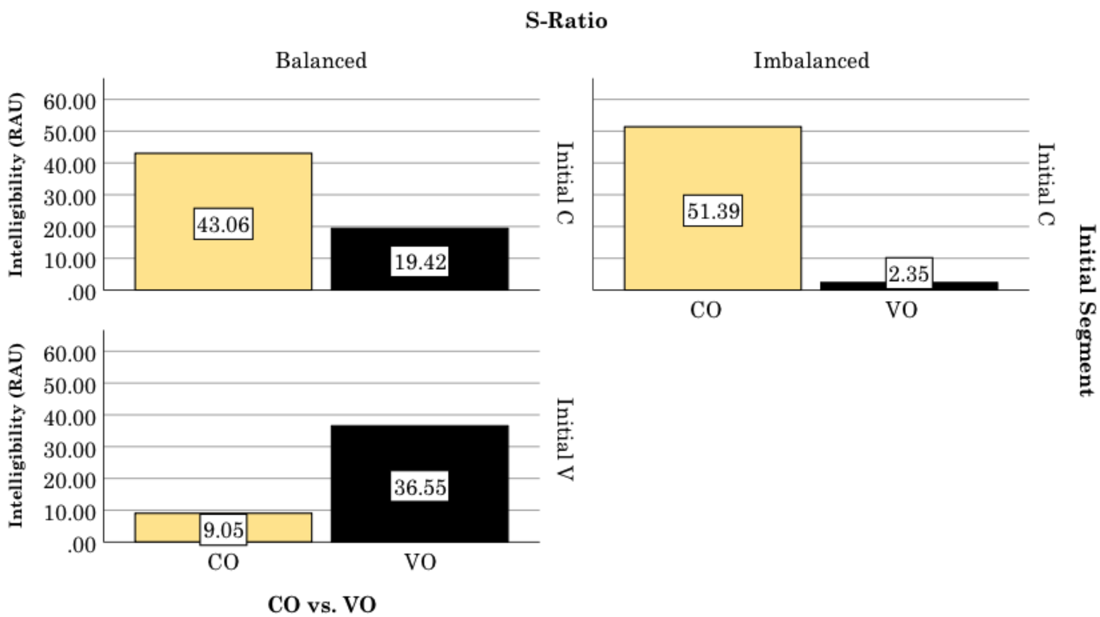
	Written Form	Spoken Form
1	6.90*	6.10
2	6.80	6.10
3	6.80	6.20
4	6.20	6.10
5	6.90	6.40
6	6.70	6.10
7	6.90	6.30
8	6.80	6.50
9	6.90	6.10
10	6.40	6.10
11	6.90	6.60
12	6.80	6.30
13	6.80	6.40
14	6.10	6.10
15	6.90	6.20
16	6.60	6.10
17	6.90	6.00
18	6.40	6.70
19	6.90	6.20

20	6.90	6.20
21	5.70	6.70
22	6.90	6.00
23	6.70	6.00
24	6.90	6.10
25	6.90	6.40
26	6.90	6.10
27	6.70	6.30
28	6.90	6.50
29	6.70	6.10
30	6.60	6.10
31	6.20	6.60
32	6.80	6.30
33	6.20	6.20
34	6.90	6.10
35	6.80	6.00
36	6.80	6.10
37	6.80	6.00
38	6.80	6.00
39	6.90	6.10
40	6.90	6.10
41	6.90	6.10

42	6.30	6.30
43	6.90	6.18
44	6.30	6.30
45	6.40	6.10
46	6.40	6.10
47	6.90	6.10
48	6.80	6.10
Mean	**6.70**	**6.21**
STD	**0.28**	**0.19**
	* Note that all values were rounded.	

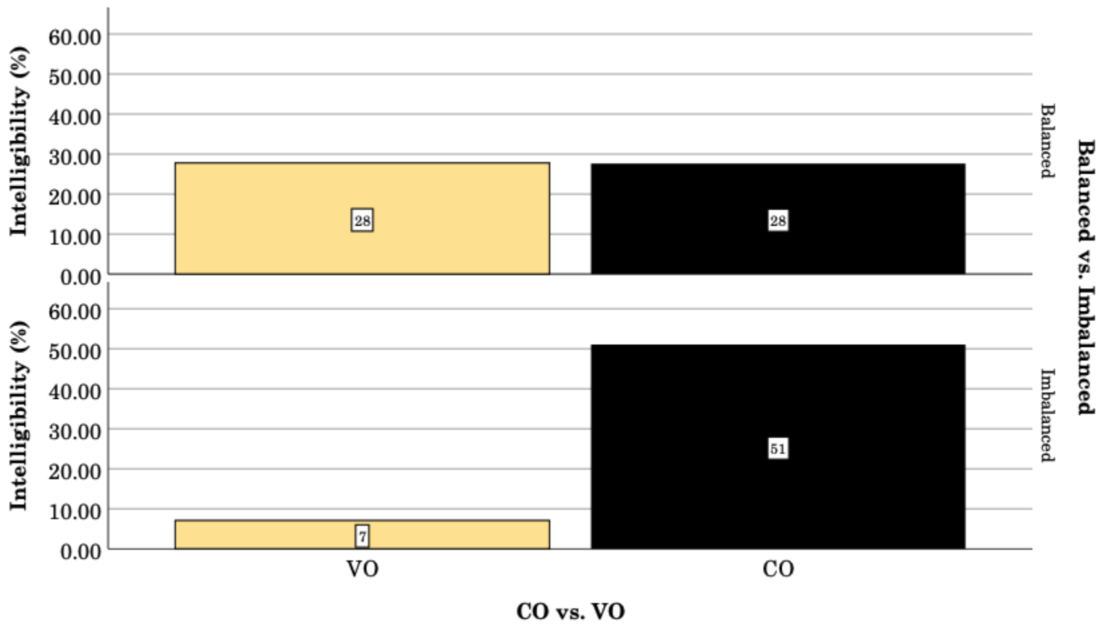
APPENDIX A2:

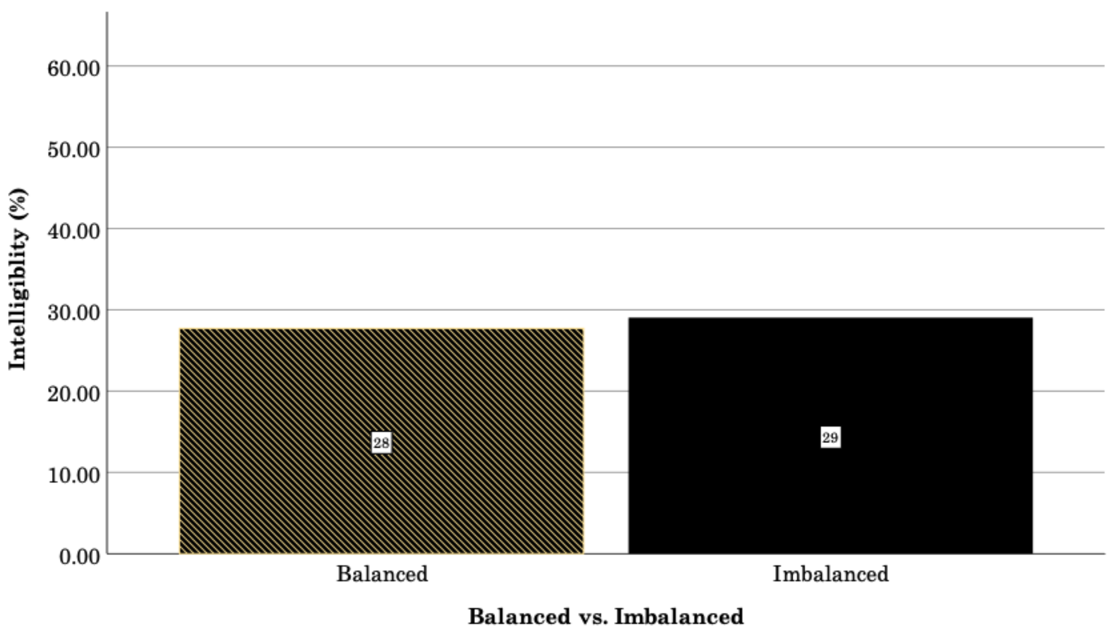
Overall Intelligibility (RAU) In Experiment 4



APPENDIX B2:

Results Grouped into Balanced vs. Imbalanced in Experiment 4





APPENDIX C2:

GEE Model Information in Experiment 4

Model information			Selection	
Dependent Variable			Intelligibility	
Probability Distribution			Binomial	
Link Function			Logit	
Subject Effect & Within-Subject/Repeated Effect			Subject & Trial	
Working Correlation Matrix Structure			Independent	
Data			**N**	
Levels	Subject Effect	Subject	20	
	Repeated Effect	Trial	48	
Subjects			20	
Measurements per Subject	Minimum		48	
	Maximum		48	
Correlation Matrix Dimension			48	
Variables and Factors			**Number**	**Percent**
Dependent Variable	Intelligibility	0	688	71.6%
		100	272	28.4%
		Total	959	100.0%
Factor	CO vs. VO	CO	480	50.0%
		VO	480	50.0%

	Total	960	100.0%
Initial Segment	Initial C	720	75.0%
	Initial V	240	25.0%
	Total	959	100.0%
S-Ratio	Balanced	480	50.0%
	Imbalanced	480	50.0%
	Total	960	100.0%

Log Quasi-Likelihood Function	Value
Quasi Likelihood Under Independence Model Criterion (QIC)	997.561
Corrected Quasi Likelihood Under Independence Model Criterion (QICC)	994.158

Observed Category	Predicted Category		Correct Percentage
	Unintelligible	Intelligible	
Unintelligible	570	117	83.0
Intelligible	150	123	44.9
Overall			72.20%

APPENDIX D2:

Complete GEE Output in Experiment 4

Parameter	B	Std. Error	95% Wald CI		Hypothesis Test			Exp(B)	95% Wald CI for Exp(B)	
			Lower	Upper	Wald χ^2	*df*	Sig.		Lower	Upper
(Intercept)	-.593	.1578	-.902	-.283	14.103	1	**.000**	.553	.406	.753
[CO vs. VO=1]	-1.487	.3034	-2.082	-.892	24.020	1	**.000**	.226	.125	.410
[CO vs. VO=0]	0[a]	.	.	.	.	.	**.**	1	.	.
[Initial Segment=1]	-.847	.2053	-1.249	-.444	17.002	1	**.000**	.429	.287	.641
[Initial Segment=0]	0[a]	.	.	.	.	.	**.**	1	.	.
[S-Ratio=1.00]	-1.135	.2147	-1.556	-.714	27.922	1	**.000**	.322	.211	.490
[S-Ratio=.00]	0[a]	.	.	.	.	.	**.**	1	.	.
[CO vs. VO=1] x [Initial Segment=1]	2.624	.3765	1.886	3.362	48.562	1	**.000**	13.789	6.592	28.843
[CO vs. VO=1] x [Initial Segment=0]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=0] x [Initial Segment=1]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=0] x [Initial Segment=0]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=1] x [S-Ratio=1.00]	1.479	.4201	.656	2.302	12.395	1	**.000**	4.388	1.926	9.996
[CO vs. VO=1] x [S-Ratio=.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=0] x [S-Ratio=1.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=0] x [S-Ratio=.00]	0[a]	.	.	.	.	.	.	1	.	.

(Scale)	1									

Dependent variable: Intelligibility. Model: (Intercept), CO vs. VO, Initial Segment, S-Ratio, CO vs. VO x Initial Segment, CO vs. VO x S-Ratio, N vs. V x S-Ratio

a. Reference category is: *Unintelligible*. b. CO is coded as 1, VO as 0, Initial C as 1, Initial V as 0, Balanced as 1, and Imbalanced as 0

			95% Wald CI		Hypothesis Test				95% Wald CI for Exp(B)	
Parameter	B	Std. Error	Lower	Up-per	Wald χ^2	*df*	Sig.	Exp(B)	Lower	Upper
(Intercept)	.593	.1578	.283	.902	14.103	1	.000	1.809	1.327	2.464
[CO vs. VO=1]	1.487	.3034	.892	2.082	24.020	1	.000	4.424	2.441	8.017
[CO vs. VO=0]	0[a]	.	.	.	.	.	.	1	.	.
[Initial Segment=1]	.847	.2053	.444	1.249	17.002	1	.000	2.332	1.559	3.487
[Initial Segment=0]	0[a]	.	.	.	.	.	.	1	.	.
[S-Ratio=1.00]	1.135	.2147	.714	1.556	27.922	1	.000	3.110	2.042	4.738
[S-Ratio=.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=1] x [Initial Segment=1]	-2.624	.3765	-3.362	-1.886	48.562	1	.000	.073	.035	.152
[CO vs. VO=1] x [Initial Segment=0]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=0] x [Initial Segment=1]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=0] x [Initial Segment=0]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=1] x [S-Ratio=1.00]	-1.479	.4201	-2.302	-.656	12.395	1	.000	.228	.100	.519
[CO vs. VO=1] x [S-Ratio=.00]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=0] x [S-Ratio=1.00]	0[a]	.	.	.	.	.	.	1	.	.

[CO vs. VO=0] x [S-Ratio=.00]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1									

Dependent variable: Intelligibility. Model: (Intercept), CO vs. VO, Initial Segment, S-Ratio, CO vs. VO x Initial Segment, CO vs. VO x S-Ratio, N vs. V x S-Ratio

a. Reference category is: *Intelligible*. b. CO is coded as 1, VO as 0, Initial C as 1, Initial V as 0, Balanced as 1, and Imbalanced as 0

APPENDIX E2:

Visual Basic Macro Code for Mean per Subject in Experiment 4

```
Sub Find_Avg()

Ws1 = "Row results"
Ws2 = "Averages"

Ws3 = "Totals"
Ws4 = "Counts"

RRHdrRow = 1
RRStrtRow = 2

RRSubjCol = 1
RRInitialCol = 2
RRSRatioCol = 3
RRCOVOCol = 4
RRIntelCol = 5

AVHdrRow = 1
AVStrtRow = 2

AVSubjCol = 2
AVStrtCol = 3

RRLastRow = Sheets(Ws1).Cells.Find(" x ", searchorder:=xlByRows,
    searchdirection:=xlPrevious).Row
```

```

Sheets(Ws2).Range("C2:T21").ClearContents
Sheets(Ws3).Cells.Clear
Sheets(Ws4).Cells.Clear

Sheets(Ws2).Cells(AVHdrRow, AVSubjCol) = "Subject"

For RRRow = RRStrtRow To RRLastRow

    RRSubject = Sheets(Ws1).Cells(RRRow, RRSubjCol)
    RRComb = Sheets(Ws1).Cells(RRRow, RRCOVOCol) & "_" & Sheets(Ws1).Cells
        (RRRow, RRInitialCol) & "_" & Sheets(Ws1).Cells(RRRow, RRSRatioCol)
    RRIntel = Sheets(Ws1).Cells(RRRow, RRIntelCol)

    AVLastRow = Sheets(Ws2).Cells.Find(" x ", searchorder:=xlByRows,
        searchdirection:=xlPrevious).Row
    AVLastCol = Sheets(Ws2).Cells.Find(" x ", searchorder:=xlByColumns,
        searchdirection:=xlPrevious).Column

    UpdRow = 0
    For AVRow = AVStrtRow To AVLastRow
        If Sheets(Ws2).Cells(AVRow, AVSubjCol) = RRSubject Then
            UpdRow = AVRow
        End If
    Next AVRow
    If UpdRow = 0 Then
        UpdRow = AVLastRow + 1
        Sheets(Ws2).Cells(UpdRow, AVSubjCol) = RRSubject
    End If

```

```vba
        UpdCol = 0
        For AVCol = AVStrtCol To AVLastCol
            If Sheets(Ws2).Cells(AVHdrRow, AVCol) = RRComb Then
                UpdCol = AVCol
            End If
        Next AVCol
        If UpdCol = 0 Then
            UpdCol = AVLastCol + 1
            Sheets(Ws2).Cells(AVHdrRow, UpdCol) = RRComb
        End If

        Sheets(Ws3).Cells(UpdRow, UpdCol) = Sheets(Ws3).Cells(UpdRow, UpdCol) + RRIntel
        Sheets(Ws4).Cells(UpdRow, UpdCol) = Sheets(Ws4).Cells(UpdRow, UpdCol) + 1
        Sheets(Ws2).Cells(UpdRow, UpdCol) = Sheets(Ws3).Cells(UpdRow, UpdCol) / Sheets(Ws4).Cells(UpdRow, UpdCol)

    Next RRRow

    ThisWorkbook.Save
    MsgBox ("Averages Recalculated!")

    End Sub

```

APPENDIX F2:

Summary of Repeated-Measures ANOVA in Experiment 4

Mauchly's Test of Sphericity

Within Subjects Effect	Mauchly's W	Approx. χ^2	*df*	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
COVO	1.000	.000	0	.	1.000	1.000	1.000
INITIAL	1.000	.000	0	.	1.000	1.000	1.000
S_RATIO	.829	3.382	2	.184	.854	.930	.500

Measure: Balanced

Source		Type III Sum of Squares	*df*	Mean Square	*F*	Sig.	Partial Eta Squared
COVO	Sphericity Assumed	33668.929	1	33668.929	31.077	.000	.621
	Greenhouse-Geisser	33668.929	1.000	33668.929	31.077	.000	.621
	Huynh-Feldt	33668.929	1.000	33668.929	31.077	.000	.621
	Lower-bound	33668.929	1.000	33668.929	31.077	.000	.621
Error(COVO)	Sphericity Assumed	20584.712	19	1083.406			
	Greenhouse-Geisser	20584.712	19.000	1083.406			
	Huynh-Feldt	20584.712	19.000	1083.406			
	Lower-bound	20584.712	19.000	1083.406			

INITIAL	Sphericity Assumed	7323.596	1	7323.596	9.593	.006	.335
	Greenhouse-Geisser	7323.596	1.000	7323.596	9.593	.006	.335
	Huynh-Feldt	7323.596	1.000	7323.596	9.593	.006	.335
	Lower-bound	7323.596	1.000	7323.596	9.593	.006	.335
Error(INITIAL)	Sphericity Assumed	14505.863	19	763.466			
	Greenhouse-Geisser	14505.863	19.000	763.466			
	Huynh-Feldt	14505.863	19.000	763.466			
	Lower-bound	14505.863	19.000	763.466			
S_RATIO	Sphericity Assumed	1748.401	2	874.201	1.547	.226	.075
	Greenhouse-Geisser	1748.401	1.708	1023.950	1.547	.229	.075
	Huynh-Feldt	1748.401	1.859	940.407	1.547	.228	.075
	Lower-bound	1748.401	1.000	1748.401	1.547	.229	.075
Error(S_RATIO)	Sphericity Assumed	21475.601	38	565.147			
	Greenhouse-Geisser	21475.601	32.443	661.956			
	Huynh-Feldt	21475.601	35.325	607.948			
	Lower-bound	21475.601	19.000	1130.295			
COVO x INITIAL	Sphericity Assumed	34.430	1	34.430	.096	.761	.005
	Greenhouse-Geisser	34.430	1.000	34.430	.096	.761	.005

	Huynh-Feldt	34.430	1.000	34.430	.096	.761	.005
	Lower-bound	34.430	1.000	34.430	.096	.761	.005
Error(COVO x INITIAL)	Sphericity Assumed	6841.685	19	360.089			
	Greenhouse-Geisser	6841.685	19.000	360.089			
	Huynh-Feldt	6841.685	19.000	360.089			
	Lower-bound	6841.685	19.000	360.089			
COVO x S_RATIO	Sphericity Assumed	8351.103	2	4175.551	5.295	.009	.218
	Greenhouse-Geisser	8351.103	1.339	6235.065	5.295	.021	.218
	Huynh-Feldt	8351.103	1.404	5949.993	5.295	.020	.218
	Lower-bound	8351.103	1.000	8351.103	5.295	.033	.218
Error(COVO x S_RATIO)	Sphericity Assumed	29964.954	38	788.551			
	Greenhouse-Geisser	29964.954	25.448	1177.490			
	Huynh-Feldt	29964.954	26.667	1123.654			
	Lower-bound	29964.954	19.000	1577.103			
INITIAL x S_RATIO	Sphericity Assumed	29464.811	2	14732.406	21.939	.000	.536
	Greenhouse-Geisser	29464.811	1.673	17617.099	21.939	.000	.536
	Huynh-Feldt	29464.811	1.815	16233.616	21.939	.000	.536
	Lower-bound	29464.811	1.000	29464.811	21.939	.000	.536
Error(INITIAL x S_RATIO)	Sphericity Assumed	25517.947	38	671.525			

	Greenhouse-Geisser	25517.947	31.778	803.014			
	Huynh-Feldt	25517.947	34.486	739.952			
	Lower-bound	25517.947	19.000	1343.050			
COVO x INITIAL x S_RATIO	Sphericity Assumed	7150.869	2	3575.434	3.308	.047	.148
	Greenhouse-Geisser	7150.869	1.835	3896.979	3.308	.052	.148
	Huynh-Feldt	7150.869	2.000	3575.434	3.308	.047	.148
	Lower-bound	7150.869	1.000	7150.869	3.308	.085	.148
Error(COVO x INITIAL x S_RATIO)	Sphericity Assumed	41070.587	38	1080.805			
	Greenhouse-Geisser	41070.587	34.865	1178.004			
	Huynh-Feldt	41070.587	38.000	1080.805			
	Lower-bound	41070.587	19.000	2161.610			

Source	COVO	INITIAL	S_RATIO	Sum of Squares	*df*	*Mean Square*	*F*	Sig.	Partial Eta Squared
COVO	Level 1 vs. Level 2			11222.976	1	*11222.976*	*31.077*	.000	.621
Error(COVO)	Level 1 vs. Level 2			6861.571	19	*361.135*			
INITIAL		Level 1 vs. Level 2		2441.199	1	*2441.199*	*9.593*	.006	.335
Error(INITIAL)		Level 1 vs. Level 2		4835.288	19	*254.489*			
S_RATIO			Level 1 vs. Level 2	503.654	1	*503.654*	*2.493*	.131	.116

			Level 2 vs. Level 3	29.697	1	*29.697*	*.075*	.787	.004
Error(S_RATIO)			Level 1 vs. Level 2	3838.954	19	*202.050*			
			Level 2 vs. Level 3	7529.482	19	*396.289*			
COVO x INITIAL	Level 1 vs. Level 2	Level 1 vs. Level 2		45.906	1	*45.906*	*.096*	.761	.005
Error(COVO x INITIAL)	Level 1 vs. Level 2	Level 1 vs. Level 2		9122.246	19	*480.118*			
COVO x S_RATIO	Level 1 vs. Level 2		Level 1 vs. Level 2	1114.176	1	*1114.176*	*.552*	.467	.028
			Level 2 vs. Level 3	8360.430	1	*8360.430*	*3.739*	.068	.164
Error(COVO x S_RATIO)	Level 1 vs. Level 2		Level 1 vs. Level 2	38359.091	19	*2018.900*			
			Level 2 vs. Level 3	42480.389	19	*2235.810*			
INITIAL x S_RATIO		Level 1 vs. Level 2	Level 1 vs. Level 2	11694.791	1	*11694.791*	*13.031*	.002	.407
			Level 2 vs. Level 3	58704.210	1	*58704.210*	*47.911*	.000	.716
Error(INITIALx S_RATIO)		Level 1 vs. Level 2	Level 1 vs. Level 2	17051.973	19	*897.472*			
			Level 2 vs. Level 3	23280.073	19	*1225.267*			
COVO x INITIAL x S_RATIO	Level 1 vs. Level 2	Level 1 vs. Level 2	Level 1 vs. Level 2	30587.100	1	*30587.100*	*3.152*	.092	.142
			Level 2 vs. Level 3	52323.359	1	*52323.359*	*5.146*	.035	.213

Error(COVO x INITIAL x S_RATIO)	Level 1 vs. Level 2	Level 1 vs. Level 2	Level 1 vs. Level 2	184368.977	19	*9703.630*			
			Level 2 vs. Level 3	193197.361	19	*10168.282*			

Measure: Imbalanced

Source		Sum of Squares	*df*	Mean Square	F	Sig.	Partial Eta Squared
COVO	Sphericity Assumed	19855.423	1	19855.423	54.358	.000	.741
	Greenhouse-Geisser	19855.423	1.000	19855.423	54.358	.000	.741
	Huynh-Feldt	19855.423	1.000	19855.423	54.358	.000	.741
	Lower-bound	19855.423	1.000	19855.423	54.358	.000	.741
Error(COVO)	Sphericity Assumed	6940.093	19	365.268			
	Greenhouse-Geisser	6940.093	19.000	365.268			
	Huynh-Feldt	6940.093	19.000	365.268			
	Lower-bound	6940.093	19.000	365.268			
S_RATIO	Sphericity Assumed	1090.547	2	545.273	1.062	.356	.053
	Greenhouse-Geisser	1090.547	1.756	621.133	1.062	.349	.053
	Huynh-Feldt	1090.547	1.920	567.894	1.062	.354	.053
	Lower-bound	1090.547	1.000	1090.547	1.062	.316	.053
Error(S_RATIO)	Sphericity Assumed	19501.694	38	513.202			
	Greenhouse-Geisser	19501.694	33.359	584.600			
	Huynh-Feldt	19501.694	36.486	534.493			

	Lower-bound	19501.694	19.000	1026.405			
COVO x S_RATIO	Sphericity Assumed	39553.306	2	19776.653	17.033	.000	.473
	Greenhouse-Geisser	39553.306	1.159	34127.592	17.033	.000	.473
	Huynh-Feldt	39553.306	1.187	33318.337	17.033	.000	.473
	Lower-bound	39553.306	1.000	39553.306	17.033	.001	.473
Error(COVO x S_RATIO)	Sphericity Assumed	44121.391	38	1161.089			
	Greenhouse-Geisser	44121.391	22.021	2003.634			
	Huynh-Feldt	44121.391	22.556	1956.123			
	Lower-bound	44121.391	19.000	2322.178			

Measure: Imbalanced

Source		Sum of Squares	*df*	Mean Square	F	Sig.	Partial Eta Squared
COVO	Sphericity Assumed	19855.423	1	19855.423	54.358	.000	.741
	Greenhouse-Geisser	19855.423	1.000	19855.423	54.358	.000	.741
	Huynh-Feldt	19855.423	1.000	19855.423	54.358	.000	.741
	Lower-bound	19855.423	1.000	19855.423	54.358	.000	.741
Error(COVO)	Sphericity Assumed	6940.093	19	365.268			
	Greenhouse-Geisser	6940.093	19.000	365.268			
	Huynh-Feldt	6940.093	19.000	365.268			
	Lower-bound	6940.093	19.000	365.268			
S_RATIO	Sphericity Assumed	1090.547	2	545.273	1.062	.356	.053

	Greenhouse-Geisser	1090.547	1.756	621.133	1.062	.349	.053
	Huynh-Feldt	1090.547	1.920	567.894	1.062	.354	.053
	Lower-bound	1090.547	1.000	1090.547	1.062	.316	.053
Error(S_RATIO)	Sphericity Assumed	19501.694	38	513.202			
	Greenhouse-Geisser	19501.694	33.359	584.600			
	Huynh-Feldt	19501.694	36.486	534.493			
	Lower-bound	19501.694	19.000	1026.405			
COVO x S_RATIO	Sphericity Assumed	39553.306	2	19776.653	17.033	.000	.473
	Greenhouse-Geisser	39553.306	1.159	34127.592	17.033	.000	.473
	Huynh-Feldt	39553.306	1.187	33318.337	17.033	.000	.473
	Lower-bound	39553.306	1.000	39553.306	17.033	.001	.473
Error(COVO x S_RATIO)	Sphericity Assumed	44121.391	38	1161.089			
	Greenhouse-Geisser	44121.391	22.021	2003.634			
	Huynh-Feldt	44121.391	22.556	1956.123			
	Lower-bound	44121.391	19.000	2322.178			

APPENDIX G2:

English Sentences with a Balanced Ratio in Experiment 5

N	Sentence
1.	Nobody saw you go to Chicago
2.	We saw the ceremony by the sea.
3.	The anonymous hacker may bully you.
4.	Lisa is an American writer.
5.	The author will follow the policy.
6.	Milwaukee is a city near Chicago.
7.	It’s rainy in Miami in summer.
8.	Andy had a coffee and a cookie.
9.	She wrote a paper on phonology.
10.	Alison was the apple of my eye.
11.	A lady met the committee today.
12.	They show a capacity to negotiate.
13.	Never borrow money to get a car.
14.	Alyssa’s daughter is at her daycare
15.	The honey bee colony may die.
16.	We decided on a leather sofa.
17.	Lay your head upon a puffy pillow
18.	Elena carried my tiny puppy
19.	Tobacco is a legal commodity
20.	Dubai is a highly dynamic city
21.	We saw a heavy military vehicle.
22.	He bought a camera for the office
23.	I decided to do psychology.
24.	Arabic is really easy to read.

APPENDIX H2:

English Sentences with an Imbalanced Ratio in Experiment 5

N	Sentence
1.	Grad students teach that class.
2.	Most students did the tasks.
3.	Grant acts like a spoiled child.
4.	Spain built new forts in March.
5.	She will leave school next spring.
6.	Cars caused hundreds of deaths.
7.	The clients got twelve tickets.
8.	Brandon's speech was not strong.
9.	Dust storms may strike Kansas.
10.	Some states have different rates.
11.	Why would one start from scratch?
12.	Large schools help small ones.
13.	Sam skipped the last three steps
14.	Grace loved her fifth-grade trip
15.	All sorts of strict rules failed
16.	Fans must name their best clubs
17.	Brooks shrugged off his stress.
18.	All funds shrunk by one sixth
19.	John struck and killed cyclists
20.	My friends don't smoke or drink
21.	Silk skirts have a waistband
22.	Most artwork looks like scrawls
23.	Read Steve's draft on French
24.	Don't let strange thoughts stick

APPENDIX I2:

Judgement Ratings in Experiment 5

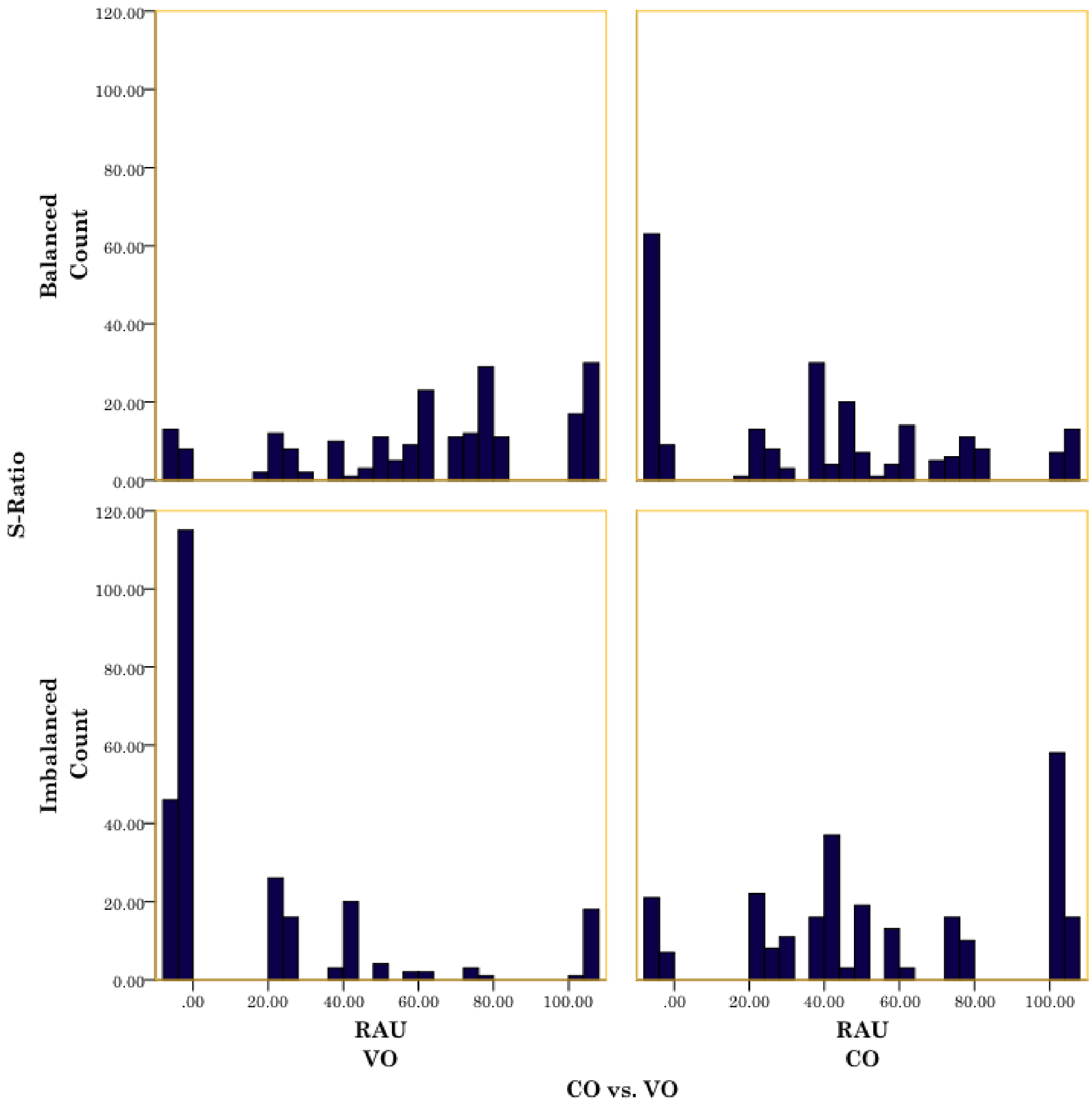
	Written Form	Spoken Form
1	6*	6.1
2	6.1	6.1
3	6	6.2
4	6.1	6.1
5	6.3	6.4
6	6	6.1
7	6.1	6.3
8	6.1	6.6
9	6	6.1
10	6	6.1
11	6.1	6.6
12	6.1	6.3
13	6.1	6.4
14	6.1	6.1
15	6.1	6.2
16	6.2	6.1
17	6.3	6
18	6.1	6.7

19	6.1	6.2
20	6	6.2
21	6	6.7
22	6	6
23	6	6
24	6.1	6.1
25	6.1	6.3
26	6	6.1
27	6	6.3
28	6.1	6.6
29	6	6.1
30	6.6	6.1
31	6.1	6.1
32	6.1	6.3
33	6.2	6.2
34	6.1	6.1
35	6.1	6
36	6.1	6.1
37	6.1	6
38	6	6
39	6.1	6.1
40	6.2	6.1

41	6.1	6.1
42	6.1	6.3
43	6.1	6.177
44	6.1	6.2
45	6.3	6.1
46	6	6.1
47	6.1	6.1
48	6.1	6.1
Mean	6.1	6.1
STD	0.10	0.18
	* Note that all values were rounded.	

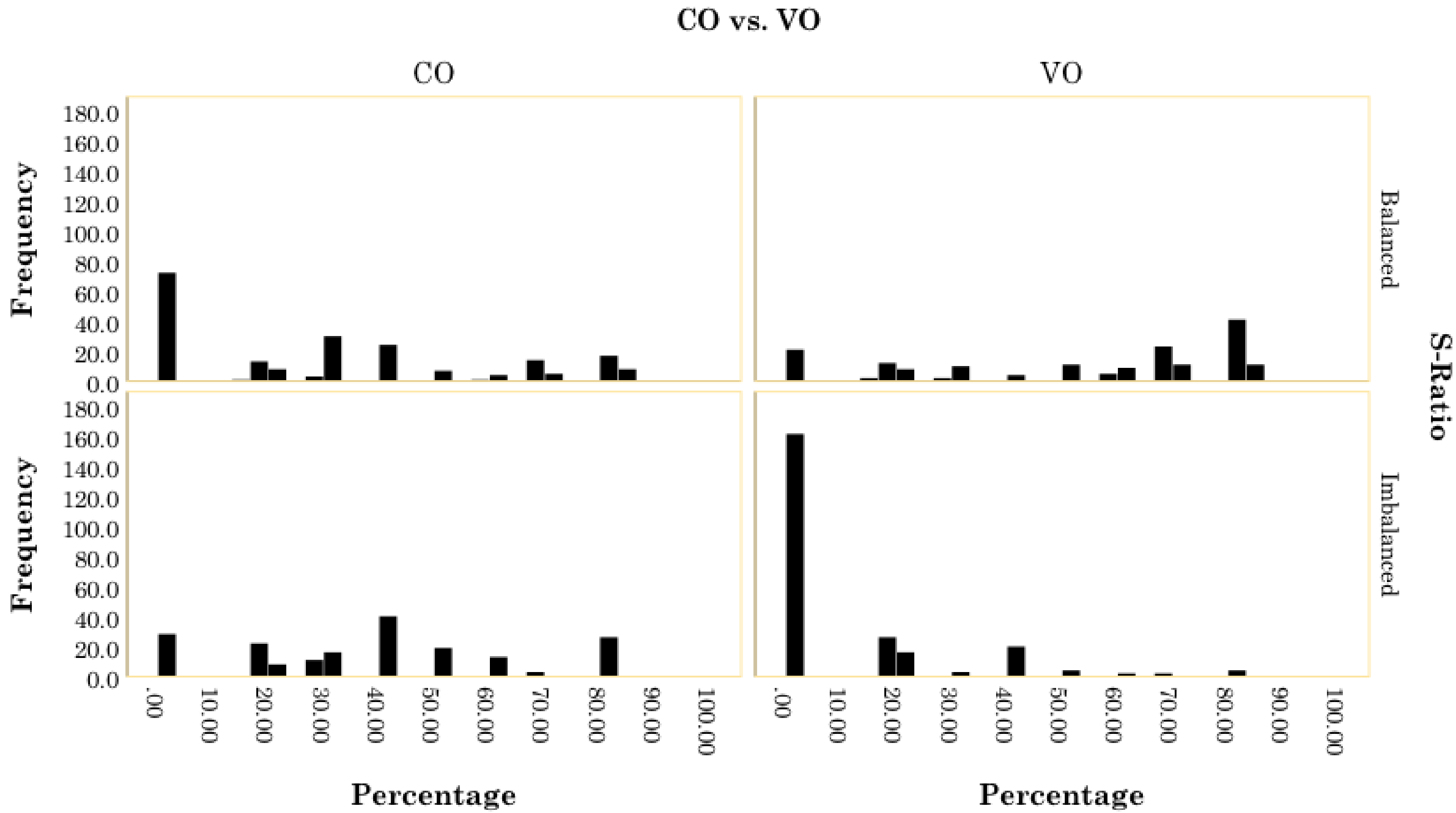
APPENDIX J2:

Frequencies of RAUs in in Experiment 5



APPENDIX K2:

Frequencies of Percentages in in Experiment 5



APPENDIX L2:

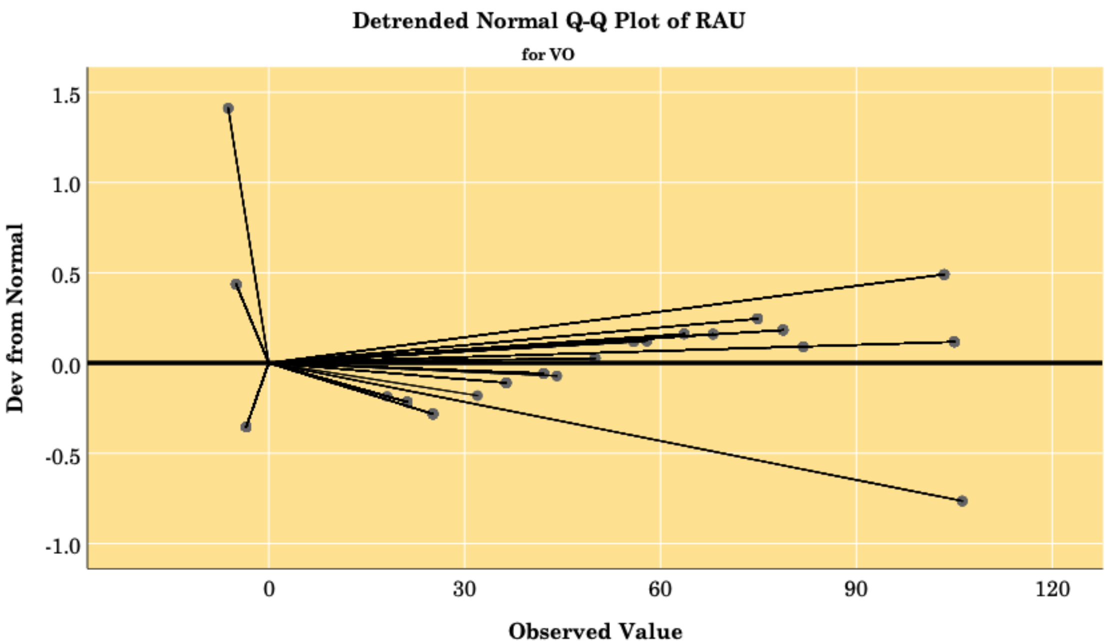
Normality Tests and Illustrations in Experiment 5

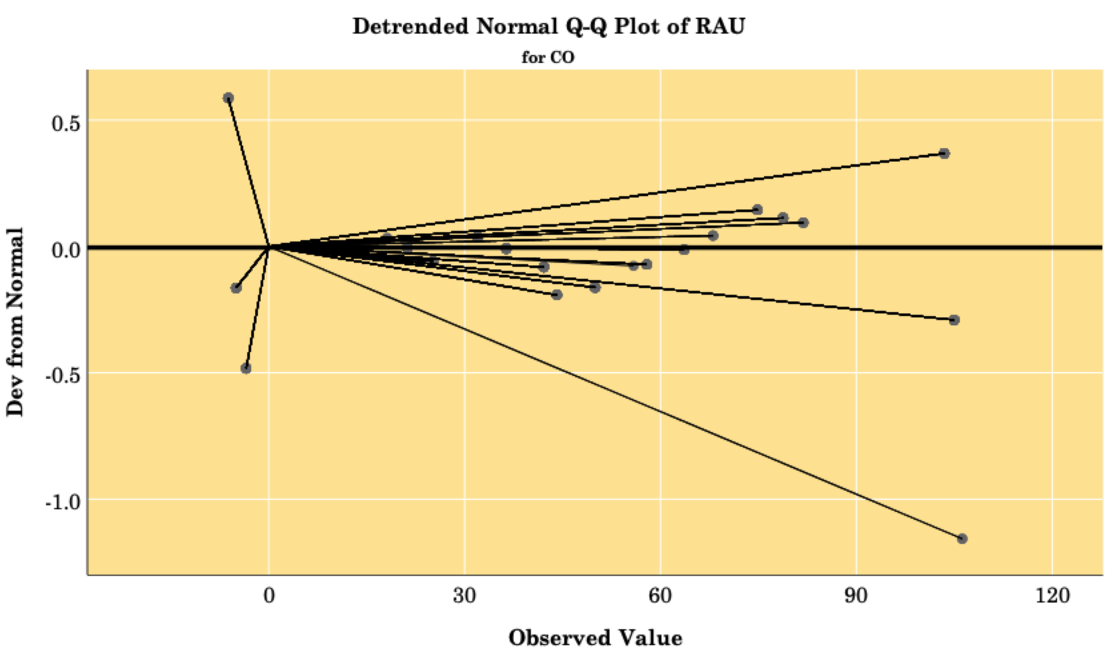
CO vs. VO		Kolmogorov-Smirnov[a] Statistic	*df*	Sig.	Shapiro-Wilk Statistic	*df*	Sig.
RAU	Balanced	.133	444	**.000**	.921	444	**.000**
	Imbalanced	.207	517	**.000**	.845	517	**.000**
	CO	.127	487	**.000**	.912	487	**.000**
	VO	.229	474	**.000**	.854	474	**.000**

Condition / Value	Skewedness	Kurtosis
Balanced	-.102	-1.128
Imbalanced	0.582	-1.004
CO	0.141	-1.081
VO	0.416	-1.256

APPENDIX M2:

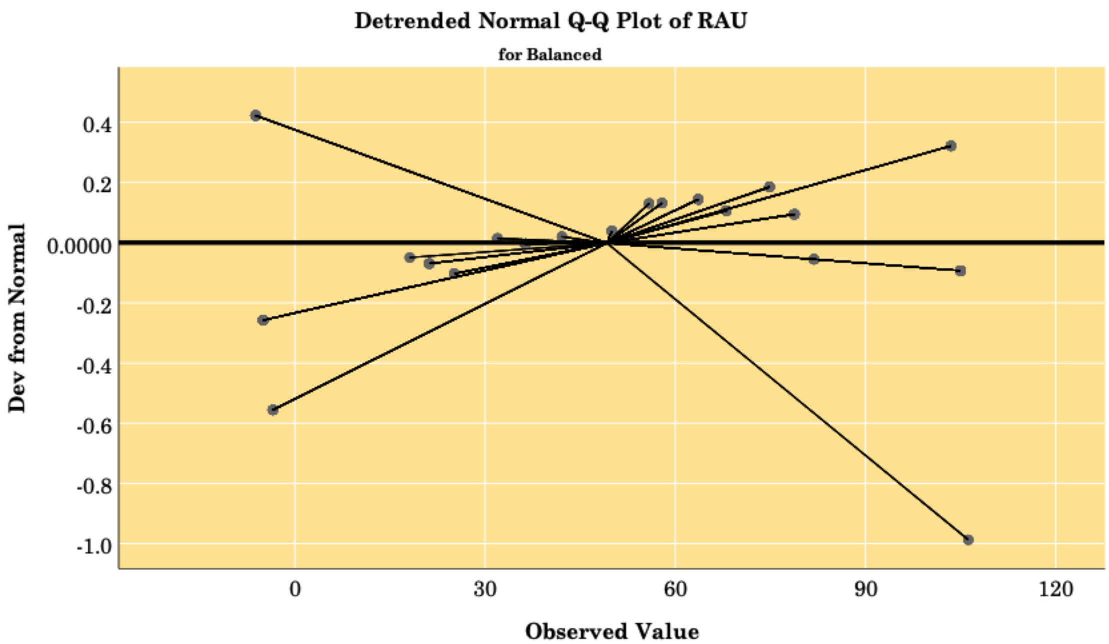
Detrended Normal Q-Q Plot of RAU for CO vs. VO

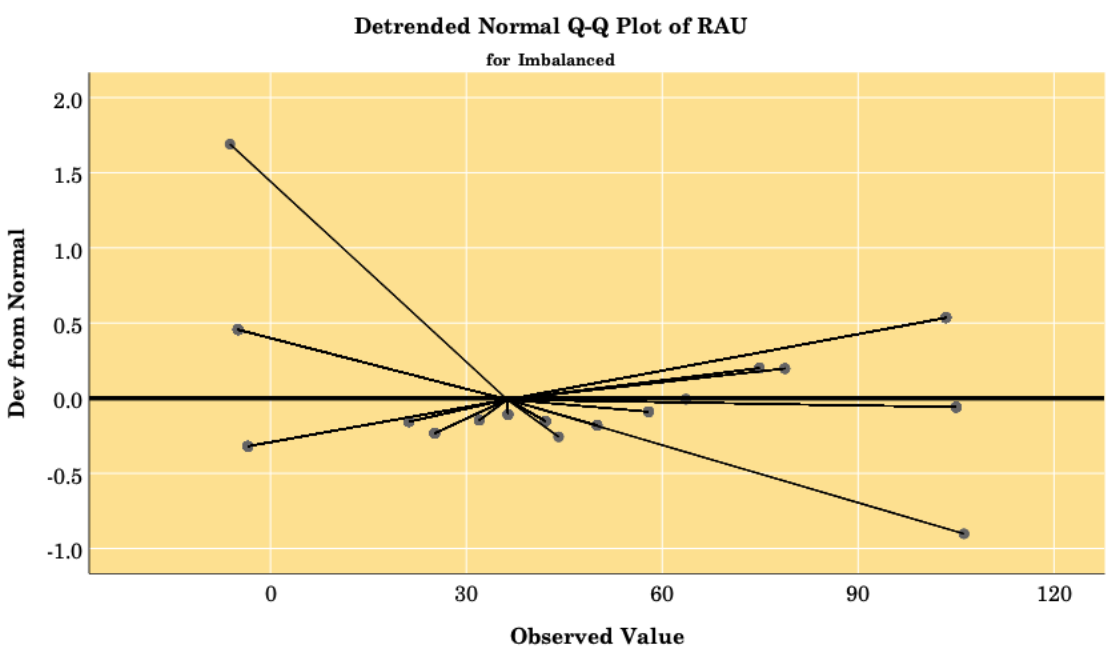




APPENDIX N2:

Detrended Normal Q-Q Plot of RAU for S-Ratio





APPENDIX O2:

GEE Model Information in Experiment 5

Model information	Selection
Dependent Variable	Intelligibility
Probability Distribution	Negative Binomial (MLE)
Link Function	Logit
Subject Effect & Within-Subject/Repeated Effect	Subject & Trial
Working Correlation Matrix Structure	Independent

Data			Numbers
Levels	Subject Effect	Subject	20
	Repeated Effect	Trial	48
Subjects			20
Measurements per Subject	Minimum		48
	Maximum		48
Correlation Matrix Dimension			48

Factor		Number	Percent
CO vs. VO	CO	480	50.0%
	VO	480	50.0%
	Total	960	100.0%
S-Ratio	Balanced	480	50.0%
	Imbalanced	480	50.0%
	Total	960	100.0%

Dependent Variable	Number	Minimum	Maximum	Mean	Std.
Intelligibility (number of words)	960	0	7	2.46	2.173

Log Quasi-Likelihood Function	Value
Quasi Likelihood Under Independence Model Criterion (QIC)[b]	7043.141
Corrected Quasi Likelihood Under Independence Model Criterion (QICC)[b]	7040.754

APPENDIX P2:

Complete GEE Output in Experiment 5

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald CI for Exp(B)	
			Lower	Upper	Wald χ^2	*df*	Sig.		Lower	Upper
(Intercept)	-.035	.0854	-.202	.133	.165	1	.685	.966	.817	1.142
[CO vs. VO=1[b]]	1.132	.0859	.964	1.300	173.662	1	.000	3.102	2.621	3.671
[CO vs. VO=0]	0[a]	.	.	.	.	.	.	1	.	.
[S-Ratio=1]	1.361	.0865	1.192	1.531	247.784	1	.000	3.901	3.293	4.621
[S-Ratio=0]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=1] x [S-Ratio=1]	-1.632	.0963	-1.821	-1.443	287.324	1	.000	.196	.162	.236
[CO vs. VO=1] x [S-Ratio=0]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=0] x [S-Ratio=1]	0[a]	.	.	.	.	.	.	1	.	.
[CO vs. VO=0] x [S-Ratio=0]	0[a]	.	.	.	.	.	.	1	.	.
(Negative Binomial)	.325[b]									

Dependent variable: Number of Words per Sentence Model: (Intercept), CO vs. VO, S-Ratio, and CO vs. VO x S-Ratio

a. CO was coded as 1, VO as 0, Balanced as 1, and Imbalanced as 0.

APPENDIX Q2:

Results from Arabic vs. English Balanced vs. Imbalanced Ratios

CO vs. VO	Arabic	English
CO	83.25%	46.50%
VO	40.44%	39.90
M	61.84	43.17

CURRICULUM VITAE

⇐ Appen-dices

Yahya Aldholmi

PERSONAL DETAILS

Yahya J. Aldholmi
Place of Birth: Faifa
Country: Saudi Arabia

College of Arts
King Saud University
Riyadh, 11451
Saudi Arabia

yaldholmi@ksu.edu.sa
&
y.aldholmi@gmail.com

TEACHING EXPERIENCE

- **Language Instructor (Internship)**
  University of Wisconsin-Milwaukee – USA, 2016
- **Lecturer**
  King Saud University –Saudi Arabia, 2011
- **Teaching Assistant**
  King Saud University –Saudi Arabia, 2007
- **Teacher**
  Private School –Saudi Arabia, 2006

QUALIFICATIONS

- **PhD in Linguistics (Experimental)**
  University of Wisconsin-Milwaukee –USA, 2018
- **MS in Educational Technology**
  Concordia University-Chicago –USA, 2018
- **MS in Computer Science- IT**
  Concordia University-Wisconsin –USA, 2018
- **Graduate Diploma in TESOL**
  University of Wisconsin-Milwaukee –USA, 2016
- **MA in Linguistics (Theoretical)**
  University of Wisconsin-Milwaukee –USA, 2014
- **MA in Linguistics (Applied)**
  University of Birmingham –UK, 2010
- **BA in Language & Literature**
  Al-Imam Moh. University, Saudi Arabia, 2005

PAPERS & CONFERENCES

- Three joint research papers on phonetics, on syntax, and on information technology are to appear in 2019.
- The 32nd **ASAL**, Arizona State University, 2018
- The 31st **ASAL**, The University of Oklahoma, 2017
- The 172nd Meeting of **ASA**, Honolulu, 2016
- The 21st **MidPhon**, Michigan State University, 2016
- The 30th **ASAL**, Stony Brook University, 2016
- The 2015 **AMP**, The University of British Columbia, 2015
- The 20th **MidPhon**, Indiana University-Bloomington, 2015