2020

Investigating the differing role of consonants and vowels in word processing

Turner, Jacqueline

http://hdl.handle.net/10026.1/16595

http://dx.doi.org/10.24382/905
University of Plymouth

Investigating the Differing Role of Consonants and Vowels in Word Processing

By

JACQUELINE TURNER

a thesis submitted to the University of Plymouth in partial fulfilment for the degree of

MASTER OF PHILOSOPHY

School of Psychology

October 2020

ACKNOWLEDGEMENTS

First and foremost, I am grateful to my supervisors, Caroline Floccia and Jeremy Goslin for their support and guidance. Thank you!

I would like to say a special thanks to Fabrice Parmentier, whose support, guidance, humour and encouragement throughout kept me going!

Thank you to Plymouth Babylab and all the children and adult participants that took part in this research. Thanks too, to the Babylab Team of researchers that I worked alongside – Claire, Hester, Sam, Paul, and the many other team members.

I would also like to acknowledge my son, Joe. Your encouragement, conversations, love and laughter, has been and continues to be, my strength. Thank you!

Finally, I would like to thank Hana, for your support and patience during some of the most difficult times in my life. Děkuji!

AUTHOR'S DECLARATION

At no time during the registration for the degree of Master of Philosophy has the author been registered for any other University award without prior agreement of the Doctoral College Quality Sub-Committee.

Work submitted for this research degree at the University of Plymouth has not formed part of any other degree either at the University of Plymouth or at another establishment.

This study was funded by a PhD studentship grant awarded from ESRC and ANR. Relevant scientific seminars and conferences were regularly attended at which work was often presented and two papers were prepared for publication. Experiment 3a in this thesis has been adapted from Durrant (2014, Unpublished doctoral dissertation).

Word count for the main body of this thesis: 40,841

Signed: Jacqueline Turner

Date: 2nd October 2020

ABSTRACT

Investigating the Differing Role of Consonants and Vowels in Word Processing

Jacqueline Turner

According to Nespor, Peña, and Mehler (2003), there is an asymmetry between consonants and vowels in language processing, with consonants being more involved at the lexical level, and vowels serving as preferential cues for grammar and prosody. The privileged role of consonants during lexical access has been demonstrated in adults across numerous languages. This consonant bias has been hypothesised to occur at the phonological level and at the earliest stages of language acquisition. The aim of this thesis is to investigate these claims in adults and toddlers.

Speech sounds fall along a sonority hierarchy (see Chapter 1), and so the contrast between consonants and vowels are not easy to define. Thus, the introduction will explain why this contrast has a special status in the study of language. Chapter 2 will trace the story from the perceptual implications found in the adult literature.

The adult experimental section (Chapter 3) investigates whether the consonant bias originates purely at the phonological level. Two identical priming experiments using both transposed and replaced stimuli were designed to tap onto either phonological or orthographical levels of processing. In experiment 1,

an auditory adaptation of the original study by Lupker, Perea and Davis (2008), we found that primes sharing consonants (e.g., BENIFET and BENAFOT) facilitated lexical access (BENEFIT) more effectively compared to primes sharing vowels (e.g., BEFENIT and BETEMIT). In experiment 2, we found the same results as Lupker et al., of no main advantage of consonants over vowels when presentation of words was visual. Our results confirmed that the consonant bias requires more than just the activation of orthographic units, as it only occurred at the phonological level. Overall, we found that the nature of the consonant bias is phonological, and the origins of the consonant transposed letter effect is orthographic.

Chapter 4 reviews the different positions around the emergence of the consonant bias. Age-related differences have been found across several languages. Thus, the infant review will trace the story from the associated developmental challenges.

The developmental experimental section (Chapter 5) explores the current ambiguity found across languages in toddlers by testing 21-month-old English toddlers. Using a preferential looking paradigm, two tasks using familiar words with either familiar distractors or unfamiliar distractors were designed to investigate differences between the perception of consonants and vowels in lexical processing. By using mispronunciations of familiar words (e.g., CAT) occurring on the onset consonant (e.g., GAT), medial vowel (CET) or coda consonant (CAD), we also looked at a potential consonantal position effect. In the first experiment which used familiar distractors, we unexpectedly found no main effect of mispronunciation. To increase sensitivity to mispronunciations, Experiment 3b used novel objects as the distractors. This time, whilst English toddlers did not exhibit a consonant bias, they appeared to make use of phonetic information incrementally.

This thesis offers a unique contribution to the consonant and vowel debate, by establishing that the consonant bias predominantly occurs at the phonological level, and that the consonant bias does not emerge for English toddlers at the onset of language acquisition.

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF TABLES

Chapter 3

Chapter 5

LIST OF FIGURES

Chapter 1

Chapter 3

LIST OF ABBREVIATIONS

C&V	Consonant and Vowel
CV hypothesis	Consonant-Vowel hypothesis
C(s)	Consonant(s)
V(s)	Vowel(s)
C-bias	Consonant bias
V-bias	Vowel bias
TL(s)	Transposed Letter(s)
RL(s)	Replaced Letter(s)
C-TL	Consonant-Transposed Letter
C-transposition(s)	Consonant transposition(s)
V-transposition(s)	Vowel transposition(s)
C-advantage	Consonant advantage
V-advantage	Vowel advantage
ALL	Artificial Language Learning
MS	milliseconds
SD	Standard Deviation
RT	Response Time
SWR	Spoken Word Recognition
CP	Correct Pronunciations
MP	Mispronunciations

Chapter 1: Introduction

Much of the information that we gather comes from either spoken or written language. From the moment we wake up, our brain begins the fascinating task of interpreting speech sounds that we perceive. Whether listening to the radio or television, reading a paper, or through face-to-face or smart device communication, we process language seamlessly. The processing of language is one of the most fundamental skills that sets humans apart from other animals. Despite our apparent ease for understanding and producing language, we host a highly complex system that includes the processing of phonetic features, phonemes, words and rules (e.g., grammar, pronouns). Language is also intimately connected to cognitive processes like perception, attention and memory, and successful communication (in most cases) depends on auditory processing. This dissertation concentrates on one part of the speech perception system that contributes to our ability to understand language – the perception of consonants and vowels in word recognition.

1.1 Consonants and vowels as universal features

All languages share the use of consonant and vowel sounds (Ladefoged & Disner, 2012). It necessarily follows that these phonological elements are a linguistic universal with a special status in the study of language (Nazzi & Cutler, 2019). Across languages, vowels tend to form the nucleus of syllables and consonants form the onset and coda. However, in some languages the nucleus can be formed by a trill syllabic-consonant. For example, in Czech the word for /ice-cream/ is /zmrzlina/ where /r/ acts as a nucleus (Gregová, 2010).

Whilst phonemes are considered to be basic speech elements, syllables have a perceptual reality (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). For instance, most of us will be able to count the number of syllables in a clearly spoken word much easier than the number of phonemes. Acoustically, the consonant and vowel contrast is not easy to define since speech sounds can be conceived of as falling along a sonority hierarchy. The sonority scale is the ranking of sounds by loudness or density (Nakajima et al., 2013)[1]. Nevertheless, whilst all languages contain phonemes at distant positions along the continuum, they all contain a distinction between vowels and consonants which is reflected in perceptual and linguistic processing (Nazzi & Cutler, 2019).

The research in this dissertation has emerged from a specific framework referred to as the Consonant-Vowel (CV) hypothesis (Nespor, Peña, & Mehler, 2003), and specifically to the predictions it generates regarding word recognition. Therefore, the aim for the following review is twofold. Firstly, we will discuss the main theoretical assumptions behind the CV hypothesis, and then present some

[1] For additional reading on the sonority in British English.

background information related to spoken word recognition, especially focusing on implemental models of word recognition.

1.2. The CV hypothesis

In a seminal paper, Nespor et al. (2003) proposed that there is a division of labour between consonants and vowels (C&V)[2], constituting the linguistic basis of the CV hypothesis. Consonants over vowels feed the lexical system and so should lead to a consonant bias (C-bias) when participants are tested with a task involving lexical processing. In contrast, vowels over consonants play a larger role in the identification of the rhythmic class of a language, together with specific properties of the syntactic structure, so should lead to a vowel bias (V-bias) when participants are tested with a task involving a type of syntactic regularity. Nespor et al. proposed that this "division of labour" might help young language learners, whereby one speech category – consonants, will help them build their lexicon, whereas vowels will be preferentially used for detecting structural regularities and so are more useful for grammatical and prosodic processing. In addition, some authors hypothesised later that infants should start processing C&Vs as distinct linguistic categories from birth thus revealing initial biases (Bonatti et al., 2005; Pons & Toro, 2010).

The CV hypothesis (Nespor et al., 2003) originated from cross-linguistic observations, including the fact that consonants are typically more numerous

[2] C&V is used as an abbreviation to 'consonant and vowel' as a singular or plural, which might also be used separately throughout this thesis.

than vowels. For example, in Malay the consonant and vowel ratio is 20C: 5V; in Arabic 29C: 3V; and in Italian 24C: 7V. Cases like Swedish with 16C: 17V; and Danish which is a highly vocalic language containing more vowels than consonants, are exceptionally rare. So, Cs generally outnumber Vs with an average of a 2.5:1 ratio cross-linguistically (Maddieson et al., 2011). Nespor et al. predicted that Cs being more numerous than Vs might be the origin of their functional specialization for lexical interpretation. However, the adult cross-linguistic evidence reveals that the preferential role of consonants for lexical distinctions goes beyond their numerical advantage and persists in languages in which there is a similar proportion of consonants and vowels (Havy, Serres, & Nazzi, 2014).

Nespor et al. (2003) also pointed out that for most languages, consonants tend to disharmonize within a word. Namely, there is a tendency for consonants which belong to the same lexical item to alternate in quality and so as a result, they become more distinctive. For example, in Arabic, adjacent root consonants produced by the same articulator is avoided (McCarthy, 1985). Vowels on the other hand, often harmonize throughout most languages. That is, since vowel harmony assimilates vowels for certain features, their distinctive power is reduced. In addition, vowel harmony is not lexical but is frequently a signal to syntax. For example, in Turkish, as well as all the affixes of a word, most of the clitics (morphemes) are syntactically attached to it, therefore signalling constituency at the lowest level. Vowels can also lose their distinctiveness independently of harmony. In nonharmonic languages like English, for instance, vowels lose distinctiveness in unstressed positions. For example, the initial vowel of the word ''about'' is a schwa (/əˈbaʊt/); in some cases the vowel can be deleted altogether, such as the second vowel in the word /sep/a/rate/.

Consonants, but not vowels often constitute morphological roots in some languages. For example, in a Semitic language such as Arabic, a classic example is the consonantal root /ktb/ which is related to write e.g., /katib/ - /writer/, /kataba/ - /he wrote/, /kitab/ - /book/. Thus, consonantal roots in Semitic languages have been an important impetus for the consonantal tier (the consonantal frame of a word), which is the level of phonological representation formed by consonants (McCarthy, 1985). Specifically, the impetus for the consonantal tier is mostly lexical. In contrast, the motivation for the vocalic tier (the information held at the vocalic level) has been of a prosodic nature. Since prosody signals syntax, Nespor et al. (2003) hypothesised that the information contained in the vocalic tier is a cue to syntax.

Consonants also tend to be produced by a temporary obstruction of the vocal tract, and vowels are produced with a relatively open vocal tract. Hence, they differ in terms of how they are perceived. A spectrogram shows that they display different patterns of acoustic energy (see Figure 1.1).

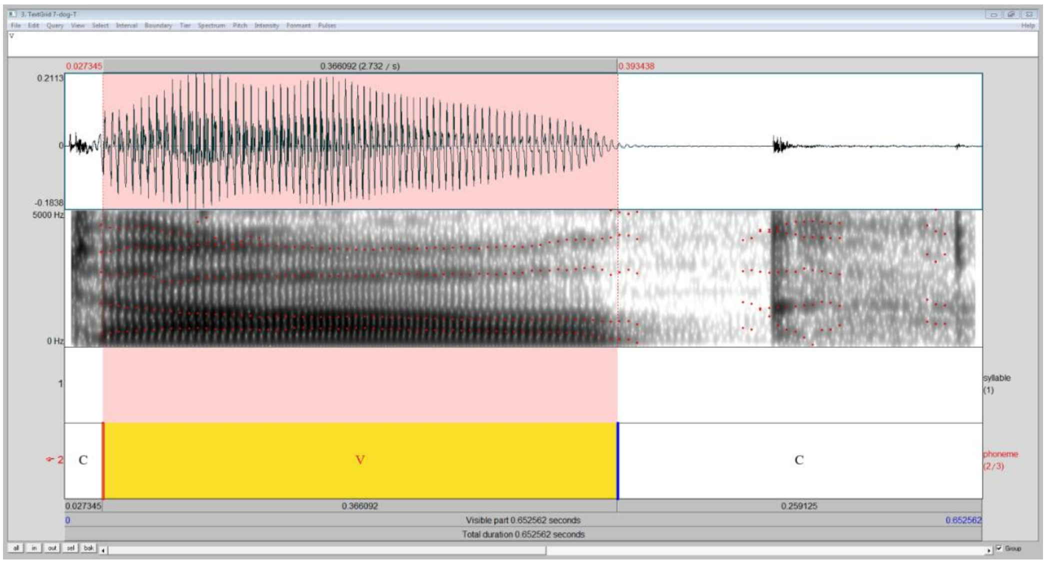

Figure 1.1: Spectrogram showing a female voice saying "dog". The highlighted area shows the vowels sound and as you can see the darker band shows that the concentration of acoustic energy is denser compared with the consonant onset and coda sounds.

Vowels are highly resonant, demonstrating at least two formant areas, making them more intense, often longer in duration, and frequently louder than consonants. Acoustic features of consonants include the notion of the formant locus, and the 'role of noise bursts' as cues to voicing and place of articulation in stops (Sussman et al., 1991). Such phenomena suggest that C&Vs are categorically distinct, independently represented in separate phonological tiers or levels (Goldsmith, 1995; McCarthy, 1988). Overall, since C&Vs play distinct roles in signalling linguistics information, Nespor et al. (2003) proposed that Cs might be more involved with lexical processing such as word identification and encoding, because they are better suited for categorical perception. Whereas, Vs are more variable, being the main carriers of prosody that marks more abstract elements, and hence provide more information about syntactic regularities.

In general, the linguistic observations discussed in this section led Nespor et al. (2003) to hypothesise because consonants are more distinctive and because most languages contain more consonants than vowels, then there might be a tendency for infants to prefer using the information provided by consonants for early word learning and early word recognition.

1.2.1 Word recognition

A classic definition for the term word recognition applies to the 'process by which perceptual input representations make contact with representations of words in the lexicon' (Pisoni & Luce, 1987). The processes associated with spoken word recognition (SWR) are broadly similar to those involved in visual word recognition (Davis, 2000). However, fundamental differences do exist between the two modalities.

Firstly, spoken language unfolds with time, calling for an incremental process. This process refers to the ability to use speech information incrementally over time. This is where an interpretation is built upon on a moment-by-moment basis from the incoming linguistic information (Swingley & Aslin, 1999). Secondly, speech is also inherently variable. For instance, variability comes from the speaker identity and gender (Coath, Brader, Fusi, & Denham, 2005), speech rate (White & Mattys, 2007), and accent or dialect (Butler, Floccia, Goslin, & Panneton, 2011). Speech is also continuous with phonological discrete gestures such as pauses and deletion of segments. In contrast, alphabetic systems are composed of distinct units (letters) formed into larger chunks (syllables, words) and sequences (sentences). Thus, the physical organisation of orthographic units is not a mirror of its spoken structure which is formed out of connected and co-articulated units (Davis, 2000). Subsequently, the fundamental differences between spoken and written word recognition have led to fundamentally different models of word recognition. Although this thesis is primarily focused on auditory word recognition, the section related to adult word recognition compares directly auditory and visual word recognition, offering a review of how current models in both modalities deal with the consonant/vowel contrast, and possibly with the claims made by the CV hypothesis.

Based on the principle of economy it has been suggested that a prelexical unit of speech acts as the interface between the acoustic signal and the lexicon (Marslen-Wilson & Warren, 1994). Traditionally various speech units have been proposed to make up these input representations (Frauenfelder & Floccia, 1999). Some theorists proposed that perceptual processing proceeds from spectral (acoustic-phonetic) representations (Klatt, 1979), distinctive phonetic features (Cornell, Lahiri, & Eulitz, 2013; Marslen-Wilson & Warren, 1994), phonemes (Pisoni & Luce, 1987), morae (Cutler & Otake, 2002), and the syllable (Mehler, Dommergues, Frauenfelder, & Segui, 1981).

Much of the reaction-time data from phoneme monitoring tasks show that syllables are processed faster than phonemes (Foss & Swinney, 1973; Segui, Frauenfelder, & Mehler, 1981). Syllables, especially in syllable-timed languages, hold a privileged position is speech perception, acquisition and production

(Goslin & Frauenfelder, 2001). Although not as reliable, syllable effects have also been found in English which is a non-syllable timed language (Mattys & Melhorn, 2005). Interestingly, it was discovered that when French listeners discriminated spoken words, results were best accounted for by advocating for the use of different types of phonological structures at temporally different stages of processing (Floccia, Kolinsky, Dodane, & Morais, 2003). The first stage specified a role for the abstract phonological structure of words (C-V frame), and a later stage involved the syllabic structure. As a result, this indicated that both syllables and distinct slots for consonants and vowels can play a role in word recognition. As will become clear in Chapter 2, numerous studies conducted in various languages reveal that consonants over vowels, hold a privilege position in adult word recognition (e.g., Delle Luche, Poltrock, Goslin, New, Floccia, & Nazzi, 2014; Lupker, Perea, & Davis., 2008; New, Araújo, & Nazzi, 2008).

1.2.2 The consonant / vowel distinction in infancy

The literature review will now present what is known about C&Vs in early words and discuss how models of word recognition account for the C&V distinction.

Despite being able to process fine-grained speech information, young babies fail just a few weeks old fail to discriminate a non-syllable-like units such as /pst/ to /tsp/, but can when the contrast occurs within a syllable-like unit e.g., /upstu/ vs. /utspu/ (Bertoncini & Mehler, 1981). Bertoncini, Bijeljac-Babic, Jusczyk, Kennedy and Mehler (1988) also found that French 2-month-olds could not detect the introduction of a new consonant syllable such as /di/ from a sequence of CV syllables, /bi/, /si/, /li/, and /mi/. However, when the syllables involved the same consonant, /bo, /ba/, /bi/, and /be/, they could detect the new vowel /bu/. This suggests that the primary perceptual unit of speech is the syllable, or vocalic nucleus (Bertoncini, Floccia, Nazzi, & Mehler, 1995; Nazzi, Bertoncini, & Mehler, 1998; van Ooijen, Bertoncini, Sansavini, & Mehler, 1997). The same conclusion was reached from studies in English which has a different stress pattern than French (Eimas, 1999; van Ooijen et al., 1997).

That said, separating French and English as belonging to two rhythmic categories, stress- and syllable-timed, is not that straight forward. Furthermore, whilst the debate around rhythmic metrics as being reliable or unreliable predictors of rhythm is interesting, it unfortunately goes beyond the scope of this thesis (e.g., see Arvanti, 2019). However, it is important to highlight some relevant differences that point towards English and French as being interesting case in terms of the consonant/vowel distinction. Whilst the French and English language share some similarities such as having the same alphabet, many variables exist that are likely to affect the phonological processing of consonants and vowels. For instance, the consonant/vowel ratio in French (17-15) and English (24-12) is not the same which should give different weight to consonantal information. Also, the English vocalic system is more complex in terms of contrastive features and diphthongs than French, which in theory should make consonants more informative in English than in French (e.g., Delle Luche et al., 2014). Interestingly, cross-linguistic developmental data indicates the exact opposite with French toddlers showing a more consistent and earlier consonant bias than English toddlers. This will be further discussed in Chapter 4.

From the age of 11 months infants start mapping word forms to meaning systematically (Gervain & Werker, 2008). The first piece of evidence examining how consonants and vowels constrain lexical access in lexical development was provided by Swingley and Aslin (2000, 2002). Using an Intermodal Preferential Looking (IPL) task (Golinkoff, Ma, Song, & Hirsh-Pasek, 2013), it was shown that English 14-, 18- to 24-month-olds reveal an equal sensitivity to mispronunciations (MP) of familiar words for vowels (/apple/ vs. /opple/ or /opal/) and consonants (/dog/ vs. /tog/ or /mog/). This finding goes against the CV hypothesis which predicts a consonant advantage in lexical processing. In addition, since the MP effect did not correlate with age or vocabulary size, Swingley et al. argued that the data does not support a developmental hypothesis for word recognition. A developmental account such as PRIMIR (Werker & Curtin, 2005), assumes that infants do not attend to phonetic details in words because it overloads their limited computational capacities. Specifically, this refers to the infant's lack of ability to be able to associate a word form to its meaning whilst at the same time as figuring out which phonetic variations are also acceptable in that word form. Instead, Swingley et al. argued for a continuity, where toddlers use their perceptual abilities as shown in earlier tasks when discriminating and categorising syllables (e.g., Bertoncini et al., 1995). The finding that infants are sensitive to consonantal MPs of words, irrespective of vocabulary size, or its position in CVC words was found to be fairly robust (Ren & Morgan, 2011; Swingley, 2009b).

Swingley (2009b) examined if English 14-month-olds process speech incrementally like adults (Dahan & Tanenhaus, 2005). Incremental processing refers to the continuous use of acoustic-phonetic information (McQueen et al., 2003). As speech unfolds competitor words consistent with the input are activated in parallel, so on hearing /gamb/ lexical contenders might include /gamble/ and /gambit/ (Fernald, Swingley, & Pinto, 2001; Swingley, Pinto, & Fernald, 1999). Specifically, Swingley (2009) investigated the phonological specificity of words with consonantal MPs in word onset and coda positions (/boat/ vs. /poat/ and /boad/) with both adults and children. It was shown that regardless of where the MP occurred, all participants fixated named targets more on hearing the correct pronunciations. Although infants were less accurate and slower than adults, both showed identical incremental temporal effects.

Additional studies have also investigated the phonological specification of numerous vowel features. For instance, it was found that English 14- and 18-months are sensitive to backness, height and roundness of vowels in familiar and learned words (Mani, Coleman, & Plunkett, 2008). This outcome indicates that vowel representations are well-specified and even small changes to some vowel features constrain lexical access. In support, another study found that Catalan-Spanish 18- to 24- month-olds also demonstrate a sensitivity to language-specific vowel contrasts such as /e/ to /ɛ/-/ ɑ/-/ i/ (Ramon-Casas et al., 2009). These studies provide evidence for an important role for vowels in lexical processing for two languages with different distributional properties in terms of consonant to vowel ratio.

In summary, the evidence reviewed at this point suggests that both consonants and vowels constrain lexical access in early lexical development which does not favour the CV hypothesis of a consonant advantage over vowels in lexical processing. However, as will become clearer in the following chapters, investigations exploring the consonant/vowel contrast is conflictual. For instance, whilst an asymmetry favouring consonants has been repeatedly observed with French toddlers (e.g. Havy et al., 2014; Nazzi, 2005; Poltrock & Nazzi, 2015) this has not been found with English infants (Floccia, Nazzi, Delle Luche, Poltrock, & Goslin, 2014; Mani & Plunkett, 2007). This thesis will explore the developmental origins of the C/V contrast and explore phoneme positional effects in early lexical representations. Furthermore, whilst the cross-linguistic adult data is less controversial showing consonantal bias effects across languages, tasks and modalities, the phonological nature of the consonant bias appears to be unclear (Delle Luche et al., 2014; Lupker et al., 2008; New et al., 2008).

This thesis will first examine whether the consonant bias effect seen in English adults (see Chapter 2) originates at the phonological level, and second whether in infants, the processing of consonants and vowels in lexical recognition is modulated by phoneme position. In what follows, we present current computational models of word recognition, to examine if and how they can account for an asymmetry in processing consonants and vowels.

1.3 Computational accounts of spoken word recognition

Various accounts of language processing typically assume initial processing stages which extract relevant perceptual information from the acoustic signal prior to later processing stages that involve lexical access (Goslin & Frauenfelder, 2001). However, the nature of the prelexical representation that are involved in each stage (Davis, 2000), including how they might or might not interact with each other, continues to be debated (e.g., Magnuson, Mirman, Luthra, Strauss, & Harris, 2018). For example, as shown in Figure 1.2, *autonomous* models such as Shortlist (Norris, 1994) and Merge (Norris, McQueen, & Cutler, 2000) can explain word recognition without the need for feedback, whereas i*nteractive* models such as TRACE (Mayor & Plunkett, 2014; McClelland & Elman, 1986) show that word recognition with feedback works better (Magnuson et al., 2018). Whilst both types of computational models represent a phonemic level of input, lexical effects are achieved through slightly different courses. The route for *autonomous* models stems from post-perceptual integration, in contrast the pathway for the *interactive* model is bidirectional.

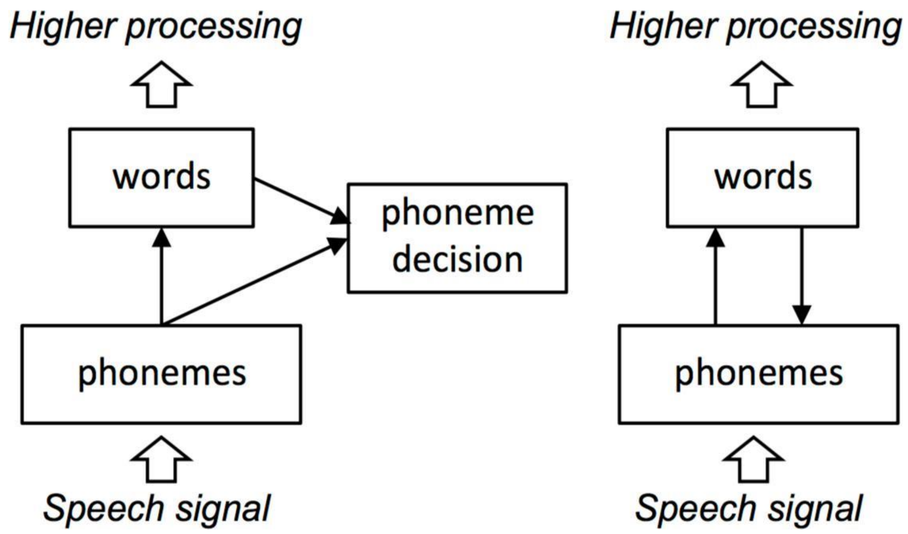

Figure 1.2: Autonomous (left) and Interactive (right) word recognition schemata (Magnuson et al., 2018).

Perhaps one of the most central feature of spoken word recognition is that the process is incremental, as opposed to visual word processing (Zwitserlood, 1989). Another dominant view is that spoken word recognition is probabilistic (Vitevitch & Luce, 1999). For instance, adult (Vitevitch et al., 1999) and infant (Gonzalez-Gomez et al., 2013) studies show that legality and probability of phonotactic patterns influence word processing and word learning.

Although the literature is replete with models of spoken word recognition (e.g. NAM: Luce & Pisoni, 1998; PARSYN: Luce, Goldinger, Auer, & Vitevitch, 2000; Cohort: Marslen-Wilson & Welsh, 1978; Shortlist: Norris, 1994), on the whole, consonant-vowel status has yet to be implemented in computational accounts of word recognition. Spoken-word recognition models do not assign any specific role to consonants over vowels, and consequently, similar priming effects are predicted. A notable exception is the highly influential TRACE model

(McClelland & Elman, 1986), which has allowed Mayor and Plunkett (2014) to replicate the consonant bias in a simulation.

Connectionist models such as TRACE operate on a parallel distributed process mapping from one representation to another through simple elements (units), sending excitatory and inhibitory signals along the way (McClelland & Elman, 1986) (see figure 1.2).

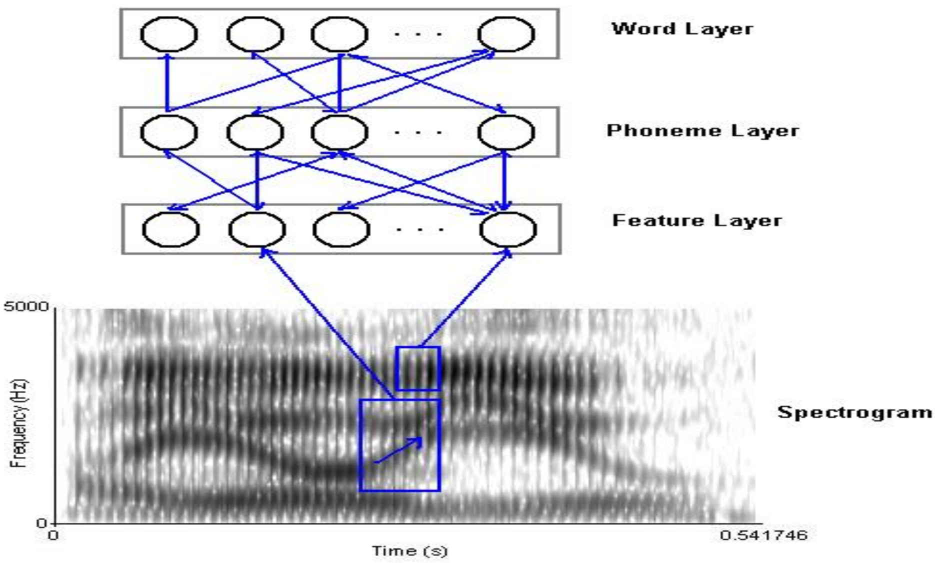

Figure 1.1: TRACE model of word recognition (e.g., McClelland & Elman, 1986).

Whilst these types of models assume that words are recognized in relation to other similar-sounding words, based on a probability of phoneme perception, they make different assumptions about which, where, and how many, lexical competitors are activated as a word unfolds. For instance, the Neighbourhood Activation Model (NAM: Luce & Pisoni, 1998) and its close relative, PARSYN (Luce et al., 2000), emphasize global similarity. Here it is predicted that words will be activated by a spoken word when they differ by no more than one phoneme (so ignoring the temporal locations of phonemes) but won't be activated if words overlap at onset (and differ by several phonemes). This means that activation is obtained for full or partial phonemic matches, but that there is no mismatch inhibition. However, priming studies have shown that phonological similarity between a prime and its target leads to lateral inhibition at the lexical level (Goldinger et al., 1989; Magnuson, Dixon, Tanenhaus, & Aslin, 2007) which is due to competition between the prime and target of real words, contrary to non-word primes where activation is contained at the phonological or pre-lexical level (Delle Luche et al., 2014). In neurobiology, lateral inhibition is the capacity of an excited neuron to reduce the activity of its neighbours. Lateral inhibition between words leads to an advantage for items overlapping at onset, due to being activated early on, they inhibit items that are activated later, such as rhymes (Magnuson et al., 2007). Cohort (Marlsen-Wilson & Welsh, 1978) and Shorlist (Norris, 1994) models on the other hand, emphasize onset-based similarity, and predict that as soon as a mismatch occurs, bottom-up inhibition takes place; these models are therefore intolerant of phonological mismatches. Along the same lines as NAM, Cohort presumes that inhibition takes place at the lexical level (albeit different mechanisms are responsible). Nonetheless, little attempt has been made to integrate the consonant bias in these models.

TRACE (McClelland & Elman, 1986) on the other hand, has been used to integrate the consonant-vowel status, in a pertinent attempt to provide a unified theoretical framework (Mayor & Plunkett, 2014). TRACE resides between models that emphasize onset similarity and those that emphasize global similarity. As well as being composed of excitatory and inhibitory nodes, these nodes are bidirectional, and unlike Shortlist, downward activation can flow from lexical to phonological nodes. Moreover, TRACE does not include explicit mismatch inhibition, and therefore similarity at any point can activate a word. In this fashion, the mechanism underlying the selection process is the same as NAM (through lateral inhibition). Even though recent findings have highlighted TRACE as being a poor predictor of adult non-word priming (see Frauenfelder et al., 2001), a recent paper evaluating TRACE's sensitivity to mispronunciations (through simulations) found that providing the parameters are set correctly (Magnuson et al., 2007; Weber & Scharenborg, 2012), TRACE can accommodate a range of infant word recognition results (Mayor et al., 2014).

Since TRACE retains a fully specified set of phonemes, meaning that consonants and vowels are coded across the same set of features, the asymmetry observed in Mayor et al. (2014) cannot be attributed to different representations or specifications for vowels and consonants. Rather, the asymmetry is claimed to arise from the increasing overrepresentation of consonants as onset phonemes relative to vowels as vocabulary size grows. Furthermore, the simulations they used involved onset consonant changes and medial vowel changes, and so the authors believed that the increased sensitivity to consonant changes is related to the increasing size of cohort competitors with vocabulary size, whereas medial vowel changes are less sensitive to changes in the number of cohort competitors.

Consequently, TRACE predicts that a language which contains a lexicon with more onset vowels than onset consonants should display an increased sensitivity to onset vowel mispronunciations, whereas sensitivity to medial consonants mispronunciations should remain stable, as observed in Danish (Højen & Nazzi, 2016).

Overall, Mayor et al. (2014) replicated the consonant-vowel asymmetry in a TRACE simulation and showed that the consonant bias arose from cohort and neighbourhood competition in a developing lexicon. Correspondingly, Delle Luche et al. (2014; see also; New & Nazzi, 2014; Soares, Perea, & Comesaña, 2014) stated that an alternative argument for the consonant bias is that phonemes do not activate exclusively on their own, but that phoneme tiers or consonantal skeletons (frames) activate the network too. This suggests that primes whose skeletons are common to few words will activate less words than primes whose frames are common to many words.

In sum, Mayor et al. (2014) have made a cogent attempt at integrating the consonant bias, by taking into account phoneme identity (consonant/vowel) with phoneme position (e.g., onsets, medial vowels) into the TRACE architecture.

Along similar lines, the consonant-vowel status has yet to be implemented in computational accounts of visual-word recognition or orthographic processing (Dunabeitia & Molinaro, 2014; Winskel & Perea, 2013). The disparity found between consonants and vowels in the visual modality poses a similar but slightly more complicated problem. To some extent this is owed to the controversy with respect to the degree to which phonology influences visual word recognition (Rastle & Brysbaert, 2006), and also partly due to the strict coding of letters asserted by archaic visual word recognition models (Yap et al., 2012). Other related issues include whether phonological codes are constructed in one single system (McClelland & Patterson, 2002), or two separate lexical and sublexical systems as in dual-route accounts (Coltheart et al., 2001), and whether these phonological codes are computed sequentially from the beginning to the end (M. Carreiras et al., 2005) or in parallel (Lee et al., 2001).

Many models of visual word processing like the interactive-activation model (Rumelhart & McClelland, 1986) and its successors, the Dual Route Cascaded model (DRC; Coltheart et al., 2001) and the Bayesian Reader model (Norris, 2006), uphold that the spatial location of each letter within a string is perfectly coded via channel specific coding. These slot type coding schemes assign separate slots for each possible letter position within a word, and letter identities are associated for each slot, thus the accurate identification of the position occurs at an absolute position. For example, the word "CAT" is represented as $C_1A_2T_3$, so the letter C is positioned in slot 1, the letter A is positioned in slot 2, and the letter T in slot 3. In contrast, the word "ACT" would be represented as $A_1C_2T_3$. Therefore, the letters C (and A) in CAT and ACT are effectively different letters (C_1 and A_2 in CAT, and C_2 and A_1 in ACT) (Kinoshita & Norris, 2013). Thus, whilst these models can explain how readers are able to distinguish anagrams such as ACT from CAT, these models fail to explain relative-position priming effects (the idea that C precedes T in the word CAT; Dunabeitia & Molinaro, 2014; Grainger, Kiyonaga, & Holcomb, 2006), and of course they therefore fail to explain the advantage of consonants over vowels for these effects. In general, interactive-activation based models have yet to accommodate differences between consonants and vowels, assuming rather that consonants and vowels are encoded in the same way and that the position of each letter within a word is perfectly encoded (Grainger, 2008).

The experimental findings that have been obtained at the level of letter position coding (transposed-letters and relative-position priming) have contributed to current models of visual-word recognition and orthographic processing, to shed their strict letter coding hypothesis in favour of noisy "slots plus slop" type of coding schemes such as the Overlap mode (Gomez et al., 2008), open-bigram models (Grainger et al., 2006; Grainger & Whitney, 2004), and spatial coding models such as the SOLAR model (Davis, 2010). Taken together, these contemporary models can accommodate such findings and are able to provide explanations for relative-position effects and transposed-letter effects. For instance, in the overlap model (Gomez et al., 2008), the identity of the letter in a string of letters is assumed to be normally distributed over position. Thus, if that string of letters happens to be the word *casino,* then the letter *a* will be associated with position 2, and also to a lesser extent (depending on the size of the standard deviation) to position 1 and 3, and even to 4 and 5. Thus, each letter which has a different standard deviation is treated as a free parameter in the model. Overall, it predicts that transposed-letter neighbours are perceptually more similar to the target word than replaced-letter neighbours. In contrast, open-bigram (OB) models propose that the order of letters in a word is coded in terms of ordered pairs (bigrams). Priming is assumed to be a function of orthographic similarity between the prime and target, calculated by the sum of OBs shared by the letter strings. For example, the word CART contains the following 6 OBs: CA, CR, CT, AR, AT, RT (bigrams are formed in correct order). The transposed-letter prime CATR shares all of the OBs except RT, i.e., it has 5 out of 6 matches. In contrast, the replaced-letter prime CABV shares with the target only one OB CA, i.e., has 1 out of 6 matches. Hence, the transposed-letter prime is more similar than the replaced-letter prime (Schoonbaert & Grainger, 2004).

In contrast, the SOLAR model (Davis 2010) postulates that order is represented as an activation gradient over all of the letters in the input, whereby the first letter has the highest activation; thereafter the level of activation progressively gets lower for each subsequent letter. For example, when transposing consonant location within a word, one simply reverses the direction of the transitional probability between those two consonants (e.g., Bonatti et al., 2005). So, in a reading model like SOLAR in which all of the letters within a word are coded independently of their position, a transposed-letter non-word such as *caniso* would activate the corresponding target word *casino*. In contrast, when consonants are replaced as in *caripo,* the computation of the transitional probabilities between consonants gives a new outcome. Bonatti et al. (2005) showed that listeners are excellent at computing transitional probabilities in consonant tiers to segment the speech stream into words, and this could explain why lexical access is not as impaired when consonants were transposed (e.g., Lupker et al., 2008) than when consonants are replaced (e.g., New et al., (2014). In addition, this could also explain why there is little difference between the transposition and replacement of vowels.

1.4 Dissertation structure

In this introduction we have outlined why investigating the role of consonants and vowels is important to word recognition, in infants and in adults. As will become clear in the following chapters, the data fuelling the consonant and vowel division of labour are mixed, which questions the nature and origin of the consonantal advantage in lexical processing. In Chapter 2, we review the cross-linguistic adult literature which has explored the consonant advantage in lexical processing in both the auditory and visual modalities. In Chapter 3, a lexical priming paradigm in both the auditory and visual domains is used to investigate the existence and the nature of the consonant bias in adults (Experiments 1 and 2). In Chapter 4, we review the developmental literature which has explored differences between consonants and vowels across various paradigms, ages and languages. In Chapter 5, we investigate how English toddlers process consonant and vowel mispronunciations in a familiar word recognition task using the inter-modal preferential looking paradigm (Experiments 3 and 4), in an attempt to reconcile some empirical differences observed in the current infant data regarding phoneme position (onset versus coda) and phoneme identity (consonant versus vowel).

Chapter 2: Introduction to the Processing of Consonants and Vowels in Adults

In Chapter 1 we discussed the contrast between consonants and vowels (C&V) and its special status as a linguistic universal (Fogerty & Humes, 2012; Ladefoged & Disner, 2012). We then presented Nespor, Peña, and Mehler's (2003) Consonant-Vowel (CV) hypothesis which proposes that Cs over Vs, are used for processing lexical information, whereas vowels are used more to compute grammar-like generalizations – the basis for this claim is explained more clearly in section 2.1.1. Essentially, the CV hypothesis focuses on the independent status of consonants and vowels which was initially based on early linguistic knowledge. Inspired by phonological and phonetic observations, as well as by neurophysiological evidence showing that Cs and Vs are processed by distinct mechanisms (Caramazza et al., 2000), Nespor et al. argued that it demonstrated a psychological reality to the C&V distinction.

As a result, numerous cross-linguistic studies have accumulated evidence of a dissociation between their functional roles showing a consonant bias (C-bias) for processing words, and vowels for processing aspects of syntax. Recent evidence in French indicates that the locus of this C-bias is at the phonological rather than the orthographical level (New & Nazzi, 2014). On this basis, the following adult experiments (Chapter 3) will compare the role of consonants and vowels in processing spoken words in English (Exp1) and in written words (Exp2), to provide a necessary cross-linguistic examination of the claim that this bias is phonological in nature.

The aim of this chapter is to review the existing cross-linguistic adult data, related to the C&V debate. Two types of studies investigating the differential role of consonants over vowels in adults have been published over the years: studies tackling auditory processing (Bonatti, Peña,, Nespor, & Mehler, 2005; Creel, Aslin, & Tanenhaus, 2006; Cutler, Sebastián-Gallés, Soler-Vilageliu, & Van Ooijen, 2000; Delle Luche et al., 2014; Toro, Nespor, Mehler, & Bonatti, 2008; Toro, Shukla, Nespor, & Endress, 2008; Van Ooijen, 1996) and studies examining visual processing (Acha & Perea, 2010; Carreiras, Gillon-Dowens, Vergara, & Perea, 2008; Duñabeitia & Carreiras, 2011; Lupker, Perea, & Davis, 2008; New, Araújo, & Nazzi, 2008; New & Nazzi, 2014). Coincidentally, most of the auditory paradigms used offline methods, whereby the participant has time to think about the language stimuli before a response is required (e.g., Cutler et al., 2000) whereas visual paradigms instead used more direct online tasks, whereby a participant processes information as it unfolds in real time and so the response to language stimuli is automatic and unconscious (e.g., New et al., 2008). Overall, most of the evidence point toward a C advantage for processing lexical information. However, as the following discussion will reveal, some of the C&V results appear to be at odds.

A key component to the CV hypothesis is that even during online processing, speech may be processed using both statistical[5] information such as frequency and co-occurrence probability (e.g., Erickson & Thiessen, 2015) and also non-statistical[6] computations, such as symbolic rule-learning (e.g., Endress & Bonatti, 2007). Consequently, statistical language acquisition refers to learning on the basis of regularities from the input (Aslin, & Newport, 1996). In contrast, non-statistical refers to learning on the basis of structural information from speech stream (Endress & Bonatti, 2007).

2.1 C&V in the Auditory Domain

The following section reviews the literature that has explored the processing of consonants and vowels in the auditory modality.

2.1.1 C&V in rules and words in adult artificial language learning

Studies using artificial languages have provided cross-linguistic evidence of a separation in the processing of consonants and vowels (e.g., Bonatti et al., 2005). Subsequently, a general agreement is that consonant frames, especially those containing initial stops, constitute the strongest regularity for word representations (e.g., Nazzi & Cutler, 2019). The following discussion will present studies that have created artificial languages exclusively designed to investigate the role of consonants in word-like representations, and the role of vowels in signalling grammatical-type representations.

[5] Statistical is probabilistic (Erickson & Thiessen, 2015).
[6] Non-statistical is symbolic (rule-learning).

Bonatti et al (2005) predicted that because in unsegmented speech, transitional probabilities [8] (TPs) are used to identify lexical segments and not used to discover grammatical-type regularities, then learners might calculate TPs on nonadjacent Cs and not Vs (Peña et al., 2002). Previously, in a seminal study, Peña et al (2002) presented learners a made-up language containing three families of trisyllabic words which were defined by nonadjacent dependencies (e.g., /puXki/, where X could be one of three syllables). Specifically, the authors called this the "AXC" language" to imply that for every item, A predicts exactly C. Thus, A_iXC_i appears with three different X's creating a family of words. For example, /puliki/, /puRaki/, /pufoki/. These were then pseudo randomised into the speech stream. The TP difference between these three families is 0.33; the TPs between the last syllable of any item and the first syllable of the following one is 0.5; and the TPs between A_i and its C_i is always .1.

In this situation, it was found that French participants could only track transitional TPs of nonadjacent CV syllables to segment a continuous stream of speech if boundary cues such as a 25ms pause were inserted between words. Segmentation cues like pauses reflect a grammatical-like regularity. Overall, Peña et al. showed that whilst adults could compute TPs to segment word-like units in unsegmented speech, they were unable to abstract rule-like regularities, which was additionally supported by neurophysiological evidence (Mueller et al., 2008). According to the CV hypothesis, this occurs because the mechanisms

[8] TPs are computations of statistical relations used to identify words and syllables (see Chapter 1).

used to support word segmentation and aspects of syntax are different. In particular, Endress and Bonatti (2007) proposed that online language processing involves two different learning mechanisms operating over one stream: a quick mechanism which extracts structural information about the stream, and a slower mechanism that detects statistical regularities among the items occurring within it.

On these grounds, Bonatti et al (2005) predicted that if consonants are more tied to word identification, learners in a word segmentation task should track TPs when implemented over Cs and not Vs. Indeed, Bonatti et al (2005) showed that French adults were able to extract families of words by calculating TPs among consonants e.g., /puragi/ /puregy/ but failed to do so when Vs carried the same statistical coherence /mopɛky/ /motɛry/. Since Bonatti et al. discovered that adults computed TPs on nonadjacent syllables, tied to the C and not V structure of the sequence, they claimed that their results were consistent with the CV hypothesis. Bonatti et al. argued that as French has an equal distribution of Cs (17) and Vs (16) and that they controlled for the token frequency of the C and V sequences, their result is not attributed to a numerical superiority between the two categories. Overall, the use of Cs seem to be privileged for lexical cues (see also Mehler et al., 2006), except in situations where the statistical computations are made simpler by allowing consecutive repetitions of the same word family (e.g., in English: Newport & Aslin, 2004). Finally, Bonatti et al. reasoned that as babies can process both C&Vs (Dehaene-Lambertz & Baillet, 1998) then linguistic constraints may emerge at the onset of language acquisition (initial bias).

However, in response to the initial bias hypothesis, Keidel, Jenison, Kluender and Seidenberg (2007) reasoned that the asymmetry found between consonants and vowels can simply be explained by the adult participants' lifetime experience with linguistics. Keidel et al. also argued that other important linguistic components such as phonology that are found to impact segmentation in online speech were not taken into account (Onnis et al., 2005). Likewise, the role of acoustic/phonetic information that lead to the development of phonological categories was not considered (Floccia et al., 2014). Instead, Keidel et al. (2007) proposed that the lexical hypothesis can account for the privileged role of Cs, whereby the participants' stored knowledge of the structure of their lexicon is the important mechanism. Thus, rather than the C-bias being driven by an innate predisposition, the lexical hypothesis prioritises the importance of lexical properties underwritten by the lexical structure and the size of the lexicon.

In response, Bonatti, Peña, Nespor and Mehler (2007) argued that structural phenomena cannot be reduced solely to statistical computations (Seidenberg et al., 2002). The main line of argument was that a series of simulations using a single-mechanism known as the Simple Recurrent Network (SRN; Elman, 1990), failed to account for 'all' of the artificial grammar learning data (Endress & Bonatti, 2007). What's more, Bonatti et al. (2007) reasoned that a lexical explanation really only applies to the first part of the CV hypothesis, that is, the lexical processing advantage for consonants. In conclusion, Bonatti et al. proposed that in word learning it is a language module which directs the statistical processor to attend to consonants over vowels.

A following study was conducted in Italian which investigated the second part of the CV hypothesis, namely, the existence of a rule mechanism (Toro, Nespor, et al., 2008). This mechanism would involve the ability to discover algebraic structures (e.g., ABB, ABA rules) over vowels and not consonants. Following Bonatti et al. (2005), they showed that when listeners are presented with tri-syllabic words created by C-sequences logical to TPs, participants could identify the words in the stream by using the distributional information. Next, they showed that when listeners were presented with vocalic sequences which followed a simple structural organisation (e.g., $V_1V_2V_1$), they were able to track a structural regularity from a vowel sequence. In a reversed test, where Vs were coherent in terms of TPs and Cs were restricted to a rule ($C_1C_2C_1$), listeners failed to use distributional information over Vs, and were unable to generalize over Cs, even with a 25ms pause between words. Toro et al. upheld that the asymmetry cannot be fully attributed to linguistic experience with mutual information (e.g., Keidel et al., 2007) as Italian has a lot more Cs (21) than Vs (7) which should have led to a greater reliance on consonants. In another study, it was found that even when Vs are made barely audible and Cs highly salient the same C&V disparity appeared (Toro, Shukla, et al., 2008). As a result, the C&V asymmetry could not be attributed to lower-level acoustic differences between C&Vs. However, Toro et al. emphasised that it is possible that acoustical and distributional differences might progressively bias the system towards the differential processing of phonological representations. To conclude, Toro et al. emphasised that the CV hypothesis predicts a switch in how participants rely on Cs and Vs, according to whether words or structural regularities are concealed in a speech stream.

In summary, the French and Italian segmentation evidence from using artificial language learning (ALL) studies support the claim that statistics are mostly performed over consonants for discovering words in a continuous stream of speech (C-bias), and that vowels are more beneficial for processing aspect of syntax. Furthermore, although Gómez et al. (2018) recently demonstrated that the role of consonants for segmenting artificial speech is also evident in Russian, tonal languages such as Cantonese and Mandarin show that vowels plus tones have a greater lexical association than consonants. In addition, whilst a comparable study in English reveals no differences between consonants and vowels (Newport & Aslin, 2004), Toro et al. (2008a) found that vowels can also be used to identify words, but only under redundant conditions. This would explain the English outcome since Newport and Aslin (2004) used immediate representations, and so non-probabilistic, of the same vocalic patterns which was also replicated in French (Bonatti et al., 2005). For example, when Bonatti et al. used immediate representation using the same vocalic sequences, the same vowel effect appeared. In contrast, when Bonatti et al. controlled for this factor by changing the structure by increasing the word families so that consecutive repetitions were avoided, vowels could not be used.

Overall, the cross-linguistic evidence confirms the role of consonants in artificial word identification. For the non-tone languages that have been tested so far, it appears that TP computations seem to be constrained at the phonemic level whereby participants compute TPs for consonants but not vowels.

2.1.2 C&V in spoken word reconstruction and word learning

In English, strong support for the different roles of consonants and vowels has come from word reconstruction tasks (Van Ooijen, 1996), and word learning tasks (Creel et al., 2006). Van Ooijen (1996) was the first to demonstrate that adults show a clear asymmetry between Cs and Vs when reconstructing words. In a free choice test, when listeners were presented with a non-word such as /kebra/ and instructed to substitute one of its phonemes to form a real word, it was found that they were more likely to make a vowel substitution e.g., /kobra/ over a consonant substitution like /zebra/. Another two conditions involved either a forced C condition or a forced V condition where participants were required to make specific phonemic changes to create a word. Here, it was found that participants made more incorrect V changes in the forced consonant condition than incorrect C changes in the forced vowel condition. Overall, this indicated that participants found the Cs more reliable than Vs when reconstructing words. English vowels are key to regional variations. For example, the Northern pronunciation of /bath/ using /æ/ rather than the Southern /ɑː/).

Consequently, Van Ooijen suggested that vowels might offer less reliability about word identity because listeners are equipped to deal with unexpected variability, particularly when there is uncertainty about a lexical candidate. To conclude, Van Ooijen showed that English listeners treat vowel identity as more mutable than consonants.

Moreover, the English data found by Van Ooijen (1996) were replicated in Dutch (Cutler et al., 2000), Japanese (Cutler & Otake, 2002) and in Spanish (Cutler et al., 2000). Additional support for word reconstruction was provided by neurophysiological data, demonstrating a stronger activation of the left inferior frontal gyrus which is an area typically involved in lexical search (Sharp et al., 2005). Taken together, the evidence showed that listeners have difficulties with making consonant substitutions which were marked by higher error rates and longer response latencies compared to the vowel substitutions. Henceforth, altogether the evidence signposted that the C&V difference in word reconstruction might be independent of the phonemic repertoire of a particular language. For example, Dutch (16V:20C) and English (17V:24C) have a relatively balanced phoneme repertoire compared to Spanish (5V:20C).

All in all, the cross-linguistic evidence from these word reconstruction tasks allowed for some previous explanations for the C&V asymmetry to be ruled-out (Nazzi & Cutler,2019). First, because all three languages vary according to the size of their C&V repertoire, interpretations based on the phoneme repertoire of a specific language were discounted. Second, since Spanish has no vowel reduction and Japanese has no lexical stress, and English contains both, a dependency on the phonology of a particular language was also rejected. Lastly, in regard to dialectal experience, especially in contrast to English, Spanish dialects vary more in Cs than in Vs, thus an explanation based on dialectal variation was also excluded. Overall, Nazzi and Cutler (2019) concluded that the results are compatible with the explanation in which participants treat non-words as two different frameworks, with the consonantal frame constraining lexical identity more than the vocalic frame.

In English, a consonant advantage has also been found in word learning tasks (Creel et al., 2006). In this study, Creel et al. showed that English listeners are more likely to confuse newly learned CVCV words like /suba/ with C-matched sequences /sabo/, than newly learned CVCV words like /diko/ with V-matched sequences like /gibo/. To explore whether Cs were more informative than Vs because of their distributional advantage, the authors switched the ratio of the segments so that Vs outnumbered Cs. Here, they found that a C-advantage still existed suggesting that simply increasing the number of Vs does not make vowels more informative. Similarly, a C-advantage was found in adults learning non-word minimal pairs in French (Havy, Serres, & Nazzi, 2014) and in Australian English (Escudero et al., 2016). Also, following the onset advantage in word learning (Magnuson et al., 2003), Creel et al. explored if Cs were privileged because of their onset position (CVCV). Because most words begin with a C or C-cluster, Creel et al. used VCVC stimuli, with the view that the C-bias should spread to later-occurring consonants in the word. Indeed, it was found that even though Cs no longer appeared in onset position, the C-bias was still present. Overall, the preference for preserving C information in syllable-onsets to select a new word reveals that consonants are more stable than vowels for lexical activation. However, a notable weakened consonant effect was observed in coda position, which is different to what was found in French where no positional effect appeared (Havy et al., 2014).

In summary, by using word-learning and word-recognition tasks the cross-linguistic evidence reveals differences between consonants and vowels in lexical processing. So far, the most conclusive outcome is that the adult C-bias is independent of the C&V phoneme repertoire of a given language. Overall, the adult C-bias appears to be a relatively robust finding in word reconstruction and word learning even in situations where the distribution of C&Vs is reversed in the inventory. Still, a somewhat inconclusive finding is the apparent positional effect in word learning found between French (Havy et al.,2014) and English (Creel et al., 2006). This outcome indicated that whilst in French the C-advantage does not appear to be modulated by the respective position of segments in a word, in English some positional modulation was found.

2.1.3 C&V in auditory primed lexical decisions

Typically, phonological priming refers to the fact that a phonological overlap between a target word and its prime results in enhanced or faster recognition of this word, as compared to an unrelated prime (Radeau et al., 1989).

Interestingly, effects of phonological overlap are not always facilitatory as initial overlaps can inhibit recognition, because of competition with alternative words that are simultaneously activated (Cutler, Van Ooijen, & Norris, 1999). In contrast, word final overlaps facilitates recognition whereby priming effects are found when the overlapping portion of the prime and target is the word's rime (Radeau et al., 1995). For example, LAMP is better primed by DAMP than by LUMP.

Using a phonological priming paradigm, Delle Luche et al. (2014) set out to explore two main goals regarding the cross-linguistic C&V debate. First, the authors aimed to clarify the role of C&Vs at the phonological level in adults by using an online measure of auditory processing. Importantly, this was undertaken because previous auditory C&V experiments had mostly been based on indirect, offline measures (e.g., Cutler et al., 2000; Van Ooijen, 1996). According to Endress and Bonatti (2007), functional differences between C&Vs emerge in online speech processing. Therefore, it was necessary to test the C&V prediction using a more direct measure. Secondly, because French (15V:17C) and English (17V:24C) differ in their consonant (and C-Cluster) and vowel ratio, and due to English having a more complex vowel system (e.g., diphthongs), the authors hypothesised that a larger C-bias should appear in English compared to French. Another interesting reason for comparing French and English adults was that the developmental literature between French and English had reported differences of when and how the C-bias emerges (see Chapter 4).

Adopting a lexical decision task which was a direct auditory adaptation of the replaced-letter (RL) paradigm used in the visual modality (New et al., 2008; New & Nazzi, 2014: see section 2.2.2), Delle Luche et al. (2014) employed non-word primes that either shared the consonants of a target word, such as /bunny/ (e.g., /benu/ – /bʌni/), the vowels (e.g., /nʌzi/ - /bʌni/), or were unrelated (e.g., /nezu/ - /bʌni/). In addition, the target words had either a VCVC or CVCV structure, and the English words were also categorized into either being trochaic or iambic (French words were considered iambic-like). For both languages they found that consonant related primes (e.g., /benu/) facilitated lexical processing for the target word (e.g., /bʌni/) compared to unrelated primes (e.g., /nezu/). Importantly, this finding agrees with the original visual task reported by New et al. (2008; 2014). However, the effect of vowel priming revealed a more complex picture. Indeed, no vocalic-priming emerged for VCVC words in both languages and in trochaic CVCV words in English which resulted in a C-bias in those conditions. However, preserving the vocalic-tier cued faster word recognition for CVCV words in French, which resulted in a vowel-bias (V-bias), and to a lesser degree in iambic CVCV English words which resulted in no bias. But further investigation revealed that rather than a vowel priming effect per se, the outcome was related to a facilitatory rhyme overlap (Radeau et al., 1995). Overall, the authors demonstrated the advantage of consonantal over vocalic information in auditory lexical processing using an online task.

To conclude, strong adult evidence for the phonological interpretation of the C-bias was provided in French and English, which is to our knowledge the only online task examining the C-advantage in the auditory modality.

2.1.4 Auditory C&V summary

In summary, the cross-linguistic evidence drawn from various auditory paradigms reveal a comparatively straightforward picture which all point to a similar conclusion. That is, they support the first part of the CV hypothesis (Nespor et al., 2004) of a greater reliance on consonants over vowels in lexical processing. In contrast, vocalic information in ALL was shown to be used to extract rule-like structures similar to that used in grammatical-like processing (e.g., Toro, Shukla, et al., 2008). Whilst an overall agreement was reached on the importance of Cs for lexical processing, the same cannot be said regarding the origins of the C-advantage. For instance, the lexical view holds that participants have learned the distribution of C and V information which prompts listeners to treat Cs and Vs differently (Keidel et al., 2007; Seidenberg et al., 2002). Another view holds that the C&V asymmetry might be modulated by some learned acoustic and/or phonological properties of language, such as the observed rhyme-bias found in English (Delle Luche et al., 2014). Lastly, whilst Endress and Bonatti (2016) maintain that learning words and rules might engage both a general and a specific-learning mechanism working in parallel, others argue for a relative all-in-one, statistical general-learning device (e.g., Laakso & Calvo, 2008, 2011; Romberg & Saffran, 2010).

Importantly, overall Delle Luche et al. (2014) demonstrated in a replaced-phoneme experiment, that when primes consisted of consonant related phonemes, a strong global consonantal priming effect on lexical decisions emerged in both English and French. As will become clear in the following section (2.2.2), this C-bias was initially found by using a visual version of the same replaced-letter task.

2.2 C&V in the Visual Domain

Within online tasks exploring lexical access, visual priming experiments also provide cross-linguistic support of a C&V asymmetry. Though, contrary to priming studies in the auditory domain, participants in these visual priming experiments are typically unaware of the prime, but the prime-target relationship influences (positively or negatively) the participants' decision about the target. For example, experiments have shown that the recognition of a written word can be facilitated by morphologically related primes (departure-DEPART) which is also independent of the targets orthographic and semantic relationship (Rastle et al., 2000). Additionally, in most languages the phonological representation between two words involve overlapping orthographical representations. Phonological words like MADE and MAID involve orthographic overlap, which makes it hard to isolate pure phonological effects (Dimitropoulou et al., 2011). Thus, the role of phonology in visual word recognition (VWR) has led to a long theoretical dichotomy (e.g., Rubenstein, Lewis, & Rubenstein, 1971; Seidenberg & McClelland, 1989).

Historically, phonological coding in VWR was thought to be processed through either assembled phonology (letter to sound correspondence), or through its orthographic structure (Frost, 1998; Jared & Seidenberg, 1991). Whilst many agreed that both phonological and visual pathways exist and work in parallel (dual-route approach), views continued to differ in terms of when and how phonology plays a role (e.g., Van Orden, Pennington, & Stone, 1990).

Weak (Coltheart et al., 2001) and strong (Berent & Perfetti, 1995) phonological views have theoretical implications regarding the mechanism(s) involved in VWR (e.g., Rastle & Brysbaert, 2006). Subsequently, the nature of cognitive processes underlining lexical decisions focuses on time course analyses, plus facilitation and inhibition to explore the involvement of (attentional) mechanisms (Neely, 1977; Perea, Moret-Tatay, & Carreiras, 2011). For instance, whilst a dual contribution involves both facilitatory and inhibitory effects (Coltheart et al., 2001; Parmentier et al., 2014), the locus of these effects occurring either on separate levels (Lee et al., 2001) or a single-route such as connectionist models (Seidenberg, 2005), remains controversial.

The following section reviews the literature that has explored the processing of consonants and vowels in the visual modality. As will become clear, although a majority of visual priming experiments converge towards a consonantal priming effect, some C-priming effects appear to be mixed.

2.2.1 C&V in written word reconstruction and word recognition

Word reconstruction tasks in the visual domain also reveal difference with how consonants and vowels are processed (Moates & Marks, 2012). For instance, by using the same word reconstruction task originally developed for the auditory modality, as discussed in section 2.1.2 (e.g., Cutler et al., 2000), the exact same results were found in the visual modality. Specifically, written words such as INVATE or HUNDLE are more likely to become INVITE or HANDLE than INVADE or BUNDLE. Overall, Moates et al. (2012) found that both English and Spanish readers produced less errors and faster responses when making consonants substitutions compared to the vowel substitutions. Thus, the consonantal frame of a word constrained written lexical identity more than its vocalic frame. Overall, the Spanish and English evidence previously found in spoken word reconstruction, was replicated across the same languages, in visual word reconstruction.

A consonant priming effect has also been observed in a backward masked priming task (Berent & Perfetti, 1995). In this paradigm, a target word is presented quickly which is then immediately followed by a non-word (backward mask). In this way, Berent et al. showed that participants identified the target word RAKE when followed by RIKK faster than when followed by RAIB. In essence, the authors found that brief durations of the C-preserving mask produced better recognition of the target word compared to brief durations of the V-preserving mask. With longer prime durations, no difference between C&V

preserving primes emerged. As a result, the authors proposed the two-cycles model of phonology assembly. The first cycle is for the phonology of consonants which is fast and automatic. The second cycle then assembles the phonology for vowels in a more controlled and slower process.

Using a delayed-letter paradigm, Lee et al. (2001) explored differences between consonants and vowels in word recognition in English. In this task, the presentation of either the consonants or vowels of a target word are delayed for 30 ms. Results from a number of priming studies using the delayed-letter paradigm, such as using T-XI or TA-I as primes preceding TAXI have revealed that the assignment of consonant labels occurs earlier than vowels in word identification (in English: Lee et al., 2001), and that consonants are more important for accessing whole-word forms (in Spanish: see Carreiras et al., 2008, for a similar finding using electroencephalographic measures). Specifically, Lee et al. (2001) hypothesised that if consonants are processed faster than vowels (Berent & Perfetti, 1995), then the cost for delaying consonants should be greater, which they measured with eye movements from a fixation point (on the target word). It was revealed that delaying the consonant presentation for 30 ms increased gaze durations of the target word relative to delaying the vowels (and an equal disruption was seen between consonants and vowels at 60 ms). Thus, the evidence was in line with Berent et al. showing that in early stages of VWR, consonant information is processed quicker than vowel information.

Relative-position priming effects also show a robust facilitatory effect for consonant primes (in Spanish: see Carreiras, Duñabeitia, & Molinaro, 2009; Duñabeitia & Carreiras, 2011, for finding using electroencephalographic measures). This effect builds on the premise that word recognition is facilitated by a prime word that has maintained the relative position of letters to the target word. For example, Duñabeitia et al. (2011) presented masked primes for a duration of 50 ms and showed that although facilitation was absent for the vowel-only primes (e.g., AIO does not prime CASINO), when primes were exclusively made up of consonants (CSN primes CASINO) a significant priming effect emerged, which replicated the findings from New et al. (2008, 2014: discussed in section 2.2.2). Nonetheless, in contrast to New et al. when the vowel or consonant primes were presented for 33 ms, they found that the asymmetry still existed.

In summary, a number of visual-word studies in English and Spanish have demonstrated that consonants are more important for accessing whole-word forms, and that consonant labels are read more rapidly than vowel labels (Lee et al., 2001). This suggest that C&V differences in visual word recognition concur towards an overall C-advantage. However, whilst these reading studies clearly support the C-bias proposal in lexical processing (Nespor et al., 2003), other visual priming studies appear to produce conflicting results.

2.2.2 C&V in replaced and transposed letter paradigms

Masked priming studies have also shown that in the early stages of VWR, phonological activation operates in isolation from orthography following quite a distinct time course in both opaque and transparent orthographies (Dimitropoulou et al., 2011). Time course analyses show that although orthographic codes are initially accessed (between 30 and 50 ms) they are quickly translated into phonological codes from 67 ms (Ferrand & Grainger, 1993), with phonological influences governing the rest of the lexical access stage (Zeguers et al., 2014).

Following the time course evidence, New et al. (2008; 2014) raised the question of whether the observed priming difference between consonants and vowels in a reading task arises at the phonological level, or at the orthographic level. The phonological interpretation would be in line with what was found with previous indirect measures (e.g., Cutler et al., 2000) and in an online measure (Delle Luche et al., 2014). Specifically, New et al. wanted to investigate a priming effect which was found for consonant (and not vowel) transpositions (Perea & Lupker, 2004). Surprisingly, this outcome first appeared to go against the proposal by Nespor et al. (2003) of a C-advantage in the lexical processing system. However, as will become clear in the following discussion, the transposed-letter priming effect most likely occurs at the orthographic level.

To explore the phonological nature of the C-advantage in lexical decisions, New et al. (2008, 2014) used replaced-letters (RLs) in a masked priming paradigm. Previous studies had established that in masked priming, phonological facilitations typically begin to emerge with prime exposures of 50 ms (declining at 67 ms), contrary to orthographic facilitation which is reliably detected around

33 ms and so before the 50 ms threshold (e.g., in French: Grainger & Ferrand, 1996; in Spanish: Pollatsek, Perea, & Carreiras, 2005). Subsequently, New et al (2008) presented forward masked primes for 50 ms. Using both VCVC and CVCV target words, they found that if readers were primed by non-words created by preserving the consonants of a target word such as DIVA (e.g., DUVO for DIVA), recognition times were faster than those preserving the vowels (RIFA). Overall, a significant disadvantage of the vowel-related prime over the unrelated prime appeared for the consonant initial and vowel initial words, revealing that the scope of the C-bias is not limited to the position of the letters/phonemes. However, since Ferrand and Grainger (1993) observed both phonological and orthographical priming effects at 50 ms, New and Nazzi (2014) conducted another study to examine if the C-bias occurs at the phonological level.

To do this, New and Nazzi (2014) replicated their earlier study except this time they manipulated the prime durations. First, they observed no priming difference between consonants and vowels with prime durations of 33 ms. Since only the orthographic code is accessed at 33 ms (e.g., Grainger et al., 1996), the result suggested that the nature of the C-bias needs more than the activation of orthographic units to emerge. Furthermore, New and Nazzi (2014) found that when they extended the prime duration to 66 ms which has been repeatedly found to activate the phonological code (Ferrand & Grainger, 1993; Zeguers et al., 2014), targets preceded by consonant-related primes were processed significantly faster than the targets preceded by vowel-related primes. Importantly, this result mirrored their first earlier study (New et al., 2008). Though, in contrast to their earlier study which established that the effect was due to C-related facilitation, the current result showed that the effect was as a result of V-related inhibition. That is, when primes were presented at 50 ms, no vowel priming was observed with the unrelated items (New et al., 2008). On the contrary, when vowel primes were presented for longer (66 ms, and at 50 ms + 16 ms) vowel priming emerged. Specifically, participants were significantly slower to respond compared to the unrelated targets, thus revealing an inhibition effect (New et al., 2014). Because their third manipulation involved the prime duration of 50 ms plus a mask for 16 ms, and the same pattern of result emerged as before when the prime duration was set at 66 ms, New and Nazzi (2014) concluded that the C-bias was not dependent on better prime consciousness.

The finding that a longer prime duration would elicit vowel activation is predicted by reading models that assign separate levels for C&Vs based on a temporal distinction (Berent et al., 1995). On the contrary, New et al. (2014) proposed that the different role of consonants and vowels in reading occurs because of an overlapping interactive activation process involving sublexical phonological influences and lexical competitors. Namely, C&V graphemes are activated at 33 ms, which corresponds to the sublexical orthographic level (Grainger & Ferrand, 1996). Then, at 50 ms activation has reached the sublexical phonological level and begins to reach the lexical level. At this stage, both sublexical and lexical influences play a role. Sublexical phonological influences are based on phoneme similarity and so are facilitative. In contrast, lexical influences are based on lateral inhibition (e.g., shared neighbourhood effects) and therefore are inhibitory (C. J. Davis & Lupker, 2006).

In conclusion, the authors argued that although different mechanisms (e.g., consonant-related facilitation and vowel-related inhibition, depending on prime duration) are responsible for the C- bias, it does not occur at the orthographical level but at the phonological level. Furthermore, the phonological interpretation of the C-bias for lexical decisions was also confirmed in an auditory replication (Delle Luche et al., 2014). Thus, the C-advantage when using a replaced-letter or replaced-phoneme experiment in masked priming does not appear to be specific to modality or to language.

Further insight into the function of consonants and vowels in visual-word recognition has transpired at the level of letter position coding, which poses additional problems for models that adopt a strict letter coding hypothesis (e.g., in English: Andrews, 1996). The transposed-letter (TL) similarity effect refers to the finding that a non-word generated by transposing letters in the middle of word is perceived highly similar as the baseword. For example, in a masked priming lexical decision task, a non-word prime containing two transposed internal letters (e.g., JUGDE) facilitates the recognition of its baseword JUDGE, more than a control prime which has been created by substituting two internal letters such as JUNPE (Perea & Lupker, 2003; Perea, Lupker, Kinoshita, & Lupker, 2003).

Using the TL paradigm, Perea and Lupker (2004) were the first to explore differences between letter identity (C&V) in lexical decisions in Spanish. In this study, they made various comparisons between contrasts (e.g., one replaced-letter primes vs. identity primes, and two replaced-letter primes vs. unrelated primes) however the critical contrast involved two nonadjacent TL letters with replacement-letters (RL) of either of the two consonants or the two vowels. Overall, they found that priming effects only appeared for C-transpositions and not for V-transpositions. That is, CANISO facilitates recognition of the target word CASINO more than a prime generated by replacing the consonants in a target word CARIVO. In support, the consonant TL (C-TL) priming effect also received electrophysiological correlates of this effect in Spanish (Carreiras, Vergara, & Perea, 2009).

Overall, C-TL priming is a robust discovery which has also been found in French (Schoonbaert & Grainger, 2004) and in English (Lupker et al., 2008) which pinpoints the importance of the consonant skeleton rather than the CV-structure in early phases of word processing (Perea, Marcet, & Acha, 2018). To conclude, the C-TL data not only reveals the brain's aptitude for flexibility with letter coding but is also the opposite to that found by New et al. (2008; 2013) and Delle Luche et al. (2014) when replacing consonants and vowels. Consequently, the C&V transposed-letter evidence does not support the hypothesis proposed by Nespor et al. (2003) of an overall C-advantage in lexical processing. That is, for the target word CASINO, the C-bias hypothesis would predict an advantage for the vowel-transposed prime CISANO because it preserves the consonantal tier, and less so for the C-TL prime CANISO which changes the C-tier as found by New et al. and Delle Luche et al. On the contrary, studies using the TL paradigm reveal a V-advantage. However, the TL paradigm has been reported to reflect orthographical processing (Perea et al., 2018; Taft, Xu, & Li, 2017) whereas the C-bias in lexical processing should only appear with phonological input. In conclusion, the C-TL effect appears to be a unique type of processing which emerges from lexical items stored in the orthographic system (Taft et al., 2017) whereby the consonantal structure of printed words is implicit to the internal lexicon (Perea et al., 2018).

In summary, a C-advantage which should only appear at the phonological level, emerges when using the RL paradigm which is not specific to modality or to language. In contrast, the TL paradigm reveals a different priming effect that goes beyond the C-bias which should only appear at the orthographic level. Thus, whilst there appears to be a transposed and replaced discrepancy in visual priming studies, differences between the TL and RL effect can be reconciled with an orthographic interpretation for the C-TL effect. The proposal that TL effects are predominantly modulated by orthographic processing and only paradigms that tap phonological rather than orthographical processes can potentially display a consonant advantage (Delle Luche et al., 2014; New et al., 2008; New & Nazzi, 2014) will be addressed in the following experimental chapter.

2.2.3 Visual C&V summary

The cross-linguistic evidence drawn from various priming experiments in the visual domain congregates towards the interpretation of a phonological based consonantal priming effect. However, whilst the C&V results from using the replaced letter paradigm are in favour of the C-bias hypothesis in lexical processing (Nespor et al., 2004), results from using the transposed letter paradigm are not so clear cut. So far, the majority of TL experiments show a facilitation with consonant TL stimuli, which is the opposite of what the consonant bias hypothesis predicts.

2.3 C&V Summary

Overall, there is converging evidence showing that consonants and vowels serve partially different roles in language, confirming the existence of a C-bias when processing spoken and written words. However, a different pattern of results appears when using replaced and transposed primes. In a replaced-phoneme paradigm, Delle Luche et al. (2014) showed a main priming effect in English when replacing the vowel segments, thus preserving the consonantal frame of the word. In a transposed-letter paradigm, Lupker et al. (2008) showed in English that transposing consonants, thus interfering with the C-skeleton of the word, leads to a priming effect, while replacing consonants does not.

However, these tasks were conducted in different modalities and so it is difficult to pinpoint where the difference comes from. For this reason, the following experiments will directly compare transposed and replaced phoneme stimuli in the auditory domain. An additional benefit of using both replaced and transposed C&Vs is that it will allow us to tap directly on phonological processing, which is where we expect the C-bias to operate. If transposed phonemes do undergo a special type of processing, beyond the consonant bias, such as temporal adjacencies and traces in working memory (Taft et al., 2017; Perea et al., 2018), then we might replicate the typical consonant-TL effect (e.g., Lupker et al., 2008), that is, an advantage for consonant-transposed primes over vowel-replaced primes.

Chapter 3: Replacing or transposing consonants and vowels in adults

The aim of the present adult study is to add to the C&V debate by clarifying the discrepancy found between transposed letter stimuli (e.g., CANISO for CASINO) found by Lupker et al. (2008) and what was found by using replaced phoneme stimuli (RIFA for DIVA) which was tested by Delle Luche et al., (2014). Essentially, in the study conducted by Delle Luche et al., (2014) they only used replaced stimuli and to make a direct comparison it would be useful to include both replaced and transposed in the same experiment. This will be achieved with an auditory adaptation of the TL paradigm used in Lupker et al. (2008) in English, for two main reasons.

Firstly, participants will be presented with spoken words only to promote phonological processing – a level of processing at which the consonant bias has been repeatedly found (Delle Luche et al., 2014; Nazzi., 2005; New & Nazzi, 2014). Second, by doing so we will disambiguate the phonological representations of the visual stimuli presented by Lupker et al. (2008). This is particularly important in English because of its opaque orthography leading to often ambiguous grapheme-to-phoneme conversions rules for phonemes, and especially for vowels (e.g., Content, 1991). For example, in the Lupker et al. stimuli, the vowel TL prime ACEDAMY for the target word ACADEMY can be produced in various ways (e.g., /acedamy/ can be produced as /əkədæmi/ or /eɪsdæmi/ or /ækɛdæmi) (e.g., Bowers et al., 2016). It is possible that in Lupker et al., the task only promoted orthographical effects because the phonological effects were masked in experimental noise, caused by these grapheme-phoneme uncertainties. By ensuring that the target word ACADEMY and all the consonant and vowel change primes will be produced as exactly as we intend them to be, we leave no room for phonological ambiguity as seen in written words.

Altogether, the argument that the consonant bias originates at the phonological level (e.g., Delle Luche et al., 2014; New et al., 2008, 2014) would receive stronger support if a different pattern emerged in an auditory version of the RL/TL experiment. Therefore, we will address this question here by using replaced and transposed consonant and vowel primes in an auditory priming study. We predict that transposing or replacing consonants will impair lexical access more than transposing or replacing vowels, revealing a phonologically based consonant bias (e.g., Delle Luche et al., 2014; Nazzi, 2005; Nazzi, et al., 2009; Nespor et al., 2003; New et al., 2008; 2014).

Experiment 1 directly investigates the existence of a consonant bias for lexical processing in the auditory modality in English. Participants were tested in an online lexical decision task in which auditory word targets such as ACADEMY (/əˈkædəmi/) were primed by auditory pseudo-words, which were created by transposing or replacing two internal nonadjacent consonants or vowels. Based on the previous findings that consonants are more relevant for lexical access than vowels, we predicted that consonant changes overall would impair word recognition more than vowel changes, an effect that was not reported in Lupker et al. (2008) presumably because of their use of phonologically ambiguous visual stimuli. Thus, we hypothesized that the reason for a lack of consonant advantage in their study was the use of phonologically ambiguous material, which would have contaminated the computation of consonant and vowel priming at the phonological level. To examine this hypothesis, we adapted the Lupker et al. experiment to the auditory modality, to (1) promote the use of phonological information and (2) control precisely for the phonological overlap between primes and targets.

The predictions regarding the distinction between transposed and replaced consonants and vowels are less straightforward. If position is key in speech processing, as is predicted by dynamic models such as Cohort (e.g., Marslen-Wilson & Welsh, 1978; Marslen-Wilson & Zwitserlood, 1989), then transposing or replacing consonants would equally impair lexical access; if, however phonemes leave a residual activation as speech unfolds, as found in the TRACE model (Mayor & Plunkett, 2012; McClelland & Elman, 1986), transposing consonants could be less disadvantageous than replacing them.

3.1 Experiment 1: Transposed and Replaced Phonemes

Method

Participants

Fifty monolingual participants, 40 females and 10 males (mean age: 22.6, SD: 3), were tested. University of Plymouth undergraduate students, together with members of the public, took part in this experiment in exchange for either course credit or a small monetary payment. All of them had normal hearing and were native speakers of English.

Stimuli

Based on the set selected by Lupker et al (2008), [9]80 English target words and their primes (including 3 new words: see below) were chosen to be between 6 and 9 characters long (mean length: 7.0), and between 2 and 5 syllables (mean: 2.9), with a mean word frequency (per million +3) count of 4.07 (SD = 0.67). This corresponds to the Zipf value of 1-3 = low frequency and 4-7 = high frequency words (subtitle UK frequency: Van Heuven et al., 2014). Following Lupker et al. (2008), for each of the target words, four pseudo-words were created to serve as primes. Also, using their exact procedure, the pseudo-words were created by: 1) transposing two nonadjacent consonants e.g., (ADACEMY-ACADEMY, the consonant transposed condition), 2) replacing those consonants with other consonants (ABANEMY-ACADEMY, the consonant replaced condition), 3)

[9] I would like to thank Colin Davis for providing the stimuli.

transposing two nonadjacent vowels (ACEDAMY-ACADEMY, the transposed vowel condition), and 4) replacing those vowels with other vowels (ACIDOMY-ACADEMY, the replaced vowel condition). All phoneme changes occurred in the middle of the word, and never in the final or initial position. See Table 3.1 for the IPA transcription of the primes for the example target word ACADEMY. The complete list of stimuli is listed in the Appendix 3A.

Table 3. 1: IPA transcription of the target word ACADEMY including the primes for each condition.

		IPA transcription
Target Word	ACADEMY	*əkædəmi*
Consonant transposed	ADACEMY	*ədækəmi*
Consonant replaced	ABANEMY	*əbænəmi*
Vowel transposed	ACEDAMY	*əkədæmi*
Vowel replaced	ACIDOMY	*əkɪdɒmi*

We ensured that the overall stress pattern and duration of the pseudo-words (e.g., ADACEMY) was controlled to be as close to the target word (e.g., ACADEMY) as possible, and that the acoustic properties for the remaining phonemes for each prime remained the same as the target word (see Table 3.1).

Because of these controls, three of the original 80 target words were replaced with three new words. For example, the original word RETIRE was replaced with DECIDE, since transposing the two nonadjacent consonants in the original word would have resulted in the articulation of a real word REWRITE.

For similar reasons, some alternative feature changes were made from the original changes in the following prime conditions for several target words. For the vowel replaced condition, 21 alternative feature changes were made, and 5 for the consonant replaced condition.

Overall, the types of feature differences consisted of single feature, two-feature, and three-feature changes (consonants: place, manner and voice) and up to four-feature changes for vowels (vowels: aperture, roundedness, place and tense). The total number of each feature change (which, again, was directly adapted from Lupker et al.) was not matched across the consonant and vowel conditions, and the consonant changes were smaller compared to the vowel changes (see Table 3.2 for the exact number of feature differences for each condition). However, this factor was weighting against our predictions that consonant changes would impair lexical access more than vowel changes.

Table 3. 2: Number of feature changes for each prime condition in the Word stimuli list.

Feature changes	**Transposed vowel**	**Replaced vowel 1**	**Replaced vowel 2**	**Transposed consonant**	**Replaced consonant 1**	**Replaced consonant 2**
1	3	8	1	24	25	31
2	51	39	48	43	33	30
3	13	16	16	13	22	19
4	13	17	15	n/a	n/a	n/a

For the non-word trials [10], again the same 80 non-words and primes used by Lupker et al (2008) were originally selected. They were between 6 and 9 phonemes long (mean length: 7.2), between 2 and 5 syllables (mean: 2.9), and the same procedure was taken as the word targets to create consonant and vowel transposed and replaced primes.

Because of similar difficulties that we encountered when adapting the word target list, fourteen of the original non-word targets were replaced because they were not pronounceable.

Although the proportion of C- and V-initial words was not balanced within each list (word and non-word list separately), they were balanced across both lists (word target = C-initial, 69; non-word target = C-initial, 68). Furthermore, unlike Lupker et al. (2008) where the average position of the first transposed/replaced letter was the same for both the vowel transpositions and for the consonant transpositions (mean = 3.1), the average position of the first transposed/replaced letter was now different for the vowel transpositions (mean = 2.9) and for the consonant transpositions (mean = 3.2). However, a partial correlation between the position of change in the word and priming score was non-significant with a coefficient of r = .060 between consonant transposed-phonemes versus vowel transposed-phonemes. Thus, any differences observed

[10] Non-word response time data were looked at for comparison, so to make it easier to follow, the term 'pseudo-word' refer to the primes and the term 'non-word' to the distractors.

between the two conditions can be attributed to the C/V differences in their identity and not to their position.

All word and non-word targets together with the primes were recorded in a soundproof booth by a female native speaker of English. All sound files within a set of stimuli were matched for duration and pitch (as summarised in Table 3.3) using Praat (Boersma & Weenink, 2011). For further details of the lexical characteristics of the target words and their primes calculated with n-watch (Davis, 2005) see the Appendix 3B. All targets were presented binaurally through headphones and were preceded by primes that came from one of the four prime conditions. The same 80-word targets and 80 non-word targets were used for all participants. To achieve the appropriate counterbalancing and avoid any repetition of the same target for a given participant, the target words were divided into four sets of 20 and each set was primed by pseudo-words from one of the four prime conditions. An identical counterbalancing procedure was taken for the non-word targets. On this basis, four lists were created which required four groups of participants to complete the counterbalancing.

Table 3.3: Mean duration times (in ms) and pitch for each condition (80 items per condition); Standard Deviations in parentheses.

Phoneme type	**Transposed**	**Transposed**	**Replaced**	**Replaced**
	Duration (ms)	Pitch	Duration (ms)	Pitch
Consonants	741.14 (94.3)	93.61 (1.3)	741.15 (93.6)	93.70 (1.3)
Vowels	740.88 (94.2)	93.17 (1.7)	741.80 (93.6)	93.43 (1.5)
Target word	Duration	Pitch		
	740.87 (93.6)	93.53 (1.4)		

Procedure

Participants were tested in individual sound-attenuated booths. Their task was to decide whether the second of two spoken words was a real word (like /SALINE/) or a pseudo-word (like /SILANE/) as quickly and as accurately as possible. The instructions were presented on the screen and they were informed that the first word they heard would always be a pseudo-word, which would be quickly followed by either a real word or pseudo-word. Using a serial response box, participants responded to real words with their dominant hand and to pseudo-words with the other hand. Ten training stimuli that did not belong to the test set were initially presented, which contained equal word and pseudo-word targets and one of the four prime conditions. Feedback about accuracy and response times was provided during the training phase. Participants were given the option to repeat the practice phase if needed and were informed that during the test phase feedback would only be provided during the mid-way break and at the very end on completion of the task.

Every participant completed 160 trials (80 experimental trials and 80 distractor trials) and were randomly assigned to one of the four lists. For example, the word ACADEMY (/*əkædəmi*/) was presented with its consonant-transposed prime /*ədækəmi*/ for one group, vowel-transposed prime /*əkədæmi*/ for the second group, consonant-replaced prime /*əbænəmi*/ for the third group and vowel-replaced prime /*əkɪdɒmi*/ for the fourth group. Each participant received a different random order of one of the four lists. The task was programmed in E-Prime version 2 (Schneider, Eschman & Zuccoloto, 2002).

Data selection

After the recording of the stimuli and the testing of participants, it was discovered that it was necessary to remove 15-word quadruplets from the analysis due to 4 unwanted vowel changes and 11 unwanted stress pattern changes[11], as well as 11 non-word quadruplets due to 5 unwanted vowel changes and 6 unwanted stress pattern changes.

For word data, the response time data to 154 incorrect responses (4.7%) were removed. Response times greater than 1402 ms (cut-off corresponding to 1% of the total trials) and responses below 250 ms were discarded (35 trials). The outlier response times that fell above and under 2.5 SD from the mean for each participant (64 responses) were excluded from the analyses (2.0%) which traditionally was the standard threshold to apply (e.g., Miller, 1991). Mean

[11] The results remained the same when these quadruplets were left in the analyses.

response times for each of the prime conditions were calculated for every participant.

For non-word data, the response time data to 284 incorrect responses (8.2% of the trials), response times greater than 1752 ms (cut-off corresponding to 1% of the trials) and responses below 250 ms were excluded from the analyses (33 trials). This was followed by removing all outliers above and less than 2.5 SD from the mean for each participant (65, that is 2.0% of the trials). Again, mean response times for each of the prime conditions were calculated for every participant.

The mean behavioural scores for each participant are typically examined in the F_1 and the F_2 analyses. This is where two analyses are conducted for response time. In the subject's analysis (F_1), condition means are obtained for every participant. For the item F_2 analyses, mean response time data were calculated for each item (65 words and 69 non-words).

Error rates were defined as the mean percentage of error for each condition for each participant (F_1) and for each item (F1), after the removal of outliers; these were out of 65 items per condition for each participant. A reduced lexical decision response time, and an increased accuracy measure predicts priming, so both (speed and accuracy) responses are informative measures of performance.

3.1.1 Results

As expected, response times for word (M = 783, SD = 79) were faster than for non-words (M = 956, SD = 99; t_1 (49) = -19.42, p = <.001; t_2 (275) = -24.24, p = <.001), and will be analysed separately.

Word Data

Repeated-measure ANOVAs based on the participant (F_1) and item (F_2) mean for correct responses were conducted with a 2 (Phoneme type: consonants, vowels) x 2 (Prime: transposition, replacement) x List [12] (List 1, List 2, List 3, and List 4) design. Phoneme and Prime type were within-participant and within-item factors, and List was a between-participant factor.

As expected, there was no main effect of List between groups in the participant data (F_1 (3, 46) = 1.20, p = .320, η_p^2 = .073), but there was a significant effect of List in the item analysis (F_2 (3, 64) = 3.92, p = .012, η_p^2 = .155). As can be seen in Figure 3.1, the tests also revealed a significant interaction between Prime and List (F_1 (3, 46) = 3.10, *p* = .036, η_p^2 = .168) and a 3-way interaction between Phoneme x Prime x List (F_1 (3, 46) = 10.94, p < .001, η_p^2 = .417). There were no significant interactions with List in the item analysis (all p >.05).

[12] List can be conceptualised as Group in the analyses and is included to extract the variance due to the error associated with the random assignment of items to lists.

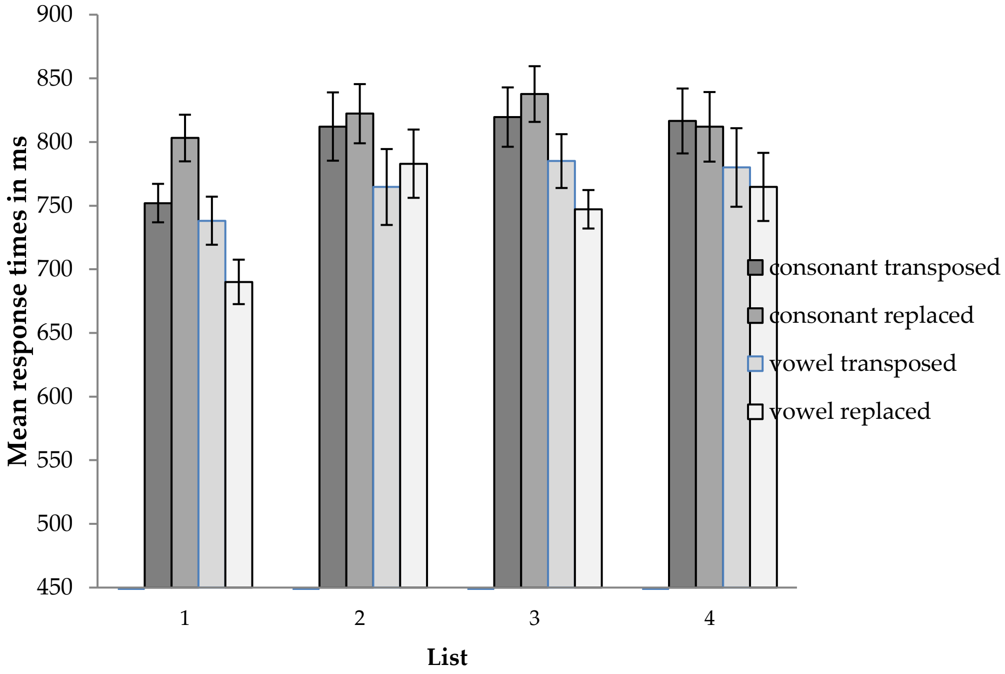

Figure 3.1: Mean response times for each list per condition. Error bars represent the standard error of the mean.

A further inspection with Prime and List indicated that for List 2, response times were significantly faster for the transposed primes (M = 788.3, SD = 96.1) than for the replaced primes (M = 802.8, SD = 91.0; t_1 (12) = 2.27, p = .042). There were no significant effects between prime conditions for the remaining three lists (all p >.05). An inspection with Phoneme and List revealed significant differences between consonants and vowels for all the Lists (all p < .01), with slower response times for consonants than vowels being found in all Lists. Planned comparisons confirmed a significant interaction between Phoneme x Prime for List 1 and List 3 (both p <.01). Given the significant interactions (F_1), List remained as a factor in the overall analyses.

As illustrated in Table 3.4, the response times that involved the vowel-modified primes were faster than response times that involved the consonant-modified primes.

Table 3. 4: Experiment 1: Mean lexical decision times (in ms) and Standard Error (in parentheses) for F_1.

Phoneme Type		**Type of Word Target**	
	Transposed	Replaced	Condition effect
Consonant	800 (*11.9*)	819 (*11.2*)	19 (-*0.7*)
Vowel	767 (*12.6*)	746 (*11.9*)	-21 (-0.7)

This was supported by a significant main effect of Phoneme type (F_1 (1, 46) = 85.16, $p < .001$, $\eta_p^2 = .649$; F_2 (1, 64) = 31.87, $p < .001$, $\eta_p^2 = .332$), corresponding to longer response times to the consonant modified words (F_1: M = 809.5, SD = 79.8, F_2 : M = 813.4, SD = 86.4) than for the vowel modified words (F_1: M = 757.1, SD = 84.3, F_2: M = 759, SD= 80.8). Indeed, replacing consonants elicited slower responses than replacing vowels (t_1 (49) = 9.61, $p <.001$; t_2 (67) = 6.20, $p <.001$), and transposing consonants also elicited slower responses than transposing vowels (t_1 (49) = 4.68, $p <.001$; t_2 (67) = 2.26, $p = .027$).

There was no main effect of Prime type (F_1 (1, 46) < 1; F_2 (1, 64) < 1). Processing targets after the transposed phoneme primes (F_1: M = 783.9, SD = 83.3, F_2: M = 787, SD= 88.2) was overall as fast as with the replaced phoneme primes (F_1: M = 783.1, SD = 77.7, F2: M = 787.1, SD= 86.3).

Finally, there was a significant interaction between these two factors (Phoneme versus Prime), (F_1: 1, 46) = 30.01, $p < .001$, $\eta_p^2 = .395$; F_2 (1, 64) = 4.41, p = .040, $\eta_p^2 = .065$). As figure 3.2 shows, listeners were significantly faster with consonant transposed primes (F_1: M = 800, SD = 84.6, F2: M = 804, SD = 89.2) than consonant replaced primes (F_1: M = 819, SD = 79.5, F_2: M = 826, SD = 80.3; t_1 (49) = 3.42, p = .001; t_2 (67) = 1.55, p = .12), whereas they were significantly slower for vowel transposed primes (F_1: M = 767, SD = 89.4, F_2: M = 770, SD = 83.5) than vowel replaced primes (F_1: M = 746, SD = 84.6, F_2: M = 749, SD = 74.9; t_1 (49) = -3.33, p = .002; t_2 (67) = -1.65, p = .10).

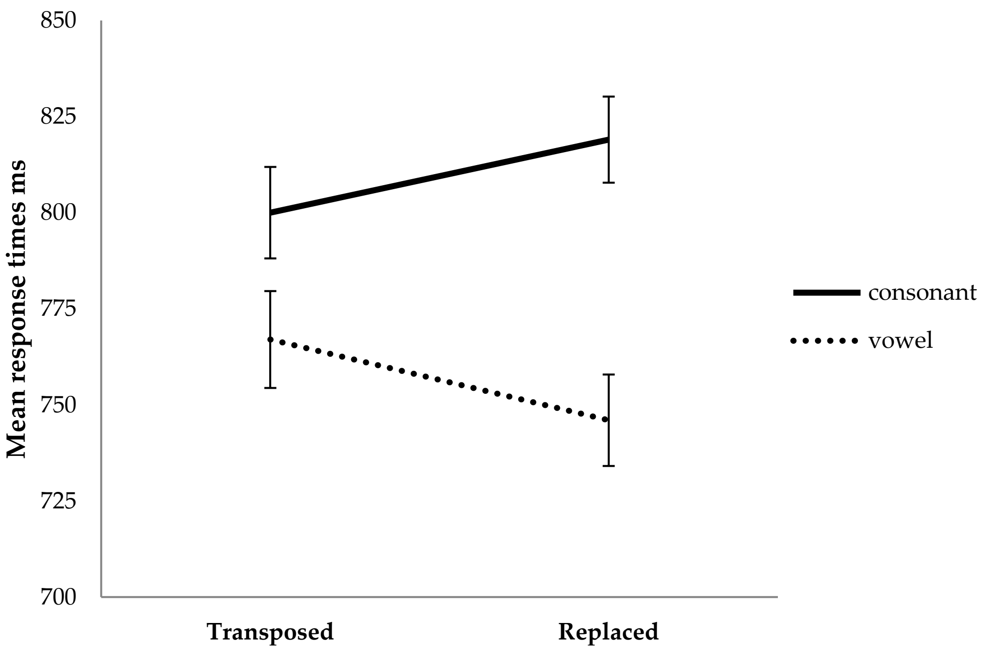

Figure 3.2: Mean response times for the four prime conditions. Error bars represent standard error of the mean.

In sum, this experiment with English listeners so far shows (1) a consonant transposed-phoneme advantage as compared to replaced consonants, and (2) a vowel replaced-phoneme advantage as compared to transposed vowels, and lastly (3) an overall advantage of vowel modifications as compared to consonant modifications.

Non-Word Data

For the non-word data, we ran again the same analysis with the 2 X 2 X 4 Phoneme X Prime X List design.

There was no effect of List between groups (F_1 (3, 46) < 1, F_2 (3, 68) < 1). However, as can be seen in Figure 3.3, a significant interaction was found between Phoneme and List (F_1 (3, 46) = 4.70, p = .006, η_p^2 = .235). There were no significant interactions with List in the item analysis (all p >.05), and none of the main effects were significant (all p > .05), so List effects in F2 analyses will not be discussed any further.

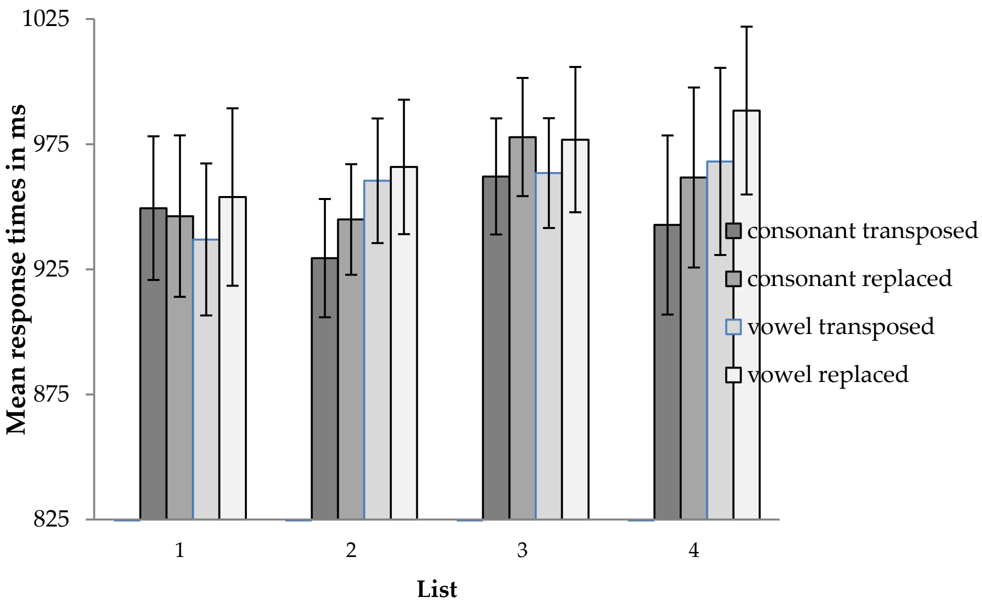

Figure 3.3: Mean response times for each list per condition. Error bars represent the standard error of the mean.

To explore the interaction between Phoneme and List in F_1 analyses, a further inspection indicated that for List 2, response times were significantly faster for the consonant primes (M = 937.1, SD = 80.7) than for the vowel primes (M = 963.1, SD = 90.6; t_1 (12) = -3.23, p = .007). Also, for List 4, response times were significantly faster for the consonant primes (M = 952.1, SD = 119.7) than for the vowel primes (M = 978.2, SD = 120; t_1 *(11)* = -3.85, p = .003). There were no other significant effects between phoneme for the remaining 2 lists (both >.05). Given the significant interactions, list remained as a factor in the overall analyses.

Contrary to the word data, there was no significant effect of Phoneme type (F_1 (1, 46) = 3.13, p = .084, η_p^2 = .064). Note that F_2 are not discussed any further due to non-significant main effects. Main results are displayed in Table 3.5.

Table 3 5: Mean lexical decision times (in ms) and Standard Error (in parentheses) for F_1 and F_2 non-word targets respectively.

Phoneme Type		**Type of Non-word Target**	
	Transposed	Replaced	Condition effect
Consonant	945 *(13.6)*	957 (*14.0*)	12 (*0.4*)
Vowel	949 (*14.1*)	971 (*15.2*)	22 (1.1)

Furthermore, there was an overall significant effect of Prime type (F_1 (1, 46) = 9.51, MSE = 1476, p = .003, η_p^2 = .171), due to longer response times for the replaced primes (M = 964.5, SD = 101.1) than for the transposed primes (M = 947.6, SD = 94.6). Post hoc comparisons revealed that this significant difference only existed between the vowel replaced condition (M = 971, SD = 107) and the vowel transposed condition (M = 949, SD = 99; t_1 (49) = 3.43, p = .001). This is different to the Word data where no main effect of Prime type was found.

More importantly, there was no significant interaction between Phoneme and Prime (F_1 (1, 46) = 1.59, p > .05, η_p^2 = .032). As figure 3.4 shows, replacing both vowels and consonants slowed down response times more than transposing these two types of phonemes.

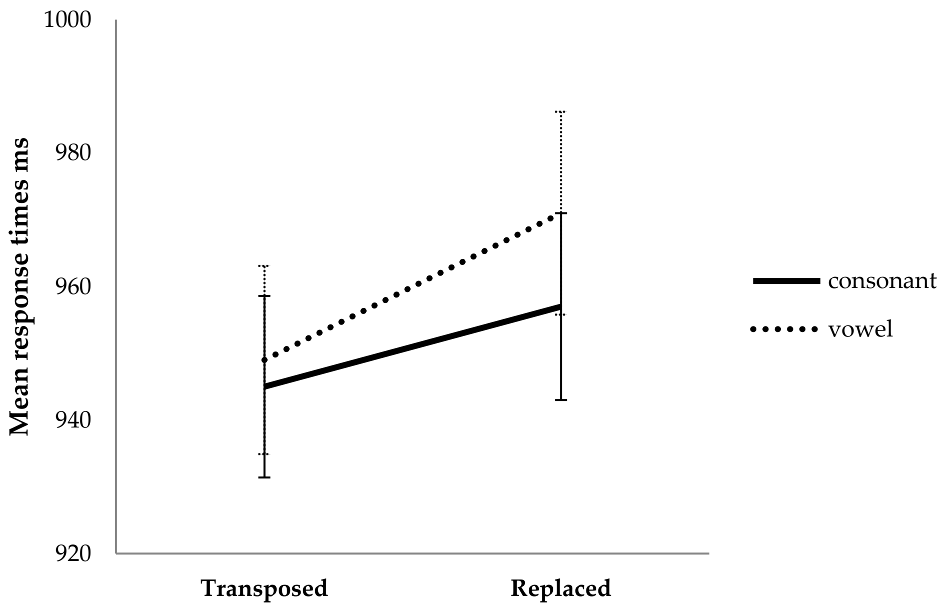

Figure 3.4: Mean response times for the non-word prime conditions. Error bars represent standard error of the mean.

To sum up, contrary to the word data, there was no main effect of Phoneme type, and no interaction between Phoneme and Prime, but a significant main effect of Prime type.

Error Rates

Error rates (for word data) were analysed with the same plan as reaction times. In brief, the error rate pattern is very similar to the RT pattern for words. This was confirmed in the main ANOVA, showing that there was a marginally significant main effect of Phoneme type (F_1 (1, 46) = 3.97, MSE = 32.70, p = .052, η_p^2 = .080). A post-hoc paired-sample t-test confirmed that more errors were made with the consonant replaced modifications (M = 6.34, SD = 7.0) than with the vowel replaced modifications (M = 3.21, SD = 4.6; t_1 (49) = 3.05, p = .004).

There was no main effect of Prime type, (F_1 (1, 46) = .718, MSE = 24.73, p > .05, η_p^2 = .003), but a significant interaction between Phoneme and Prime (F_1 (1, 46) = 7.52, MSE = 15.10, p < .01, p = .009, η_p^2 = .141). Listeners made more errors with consonant replaced primes than consonant transposed primes whereas they made fewer errors for vowel replaced primes compared with the vowel transposed primes. Figure 3.5 clearly shows that the phoneme type influenced accuracy performance, and that this was contingent on the type of prime.

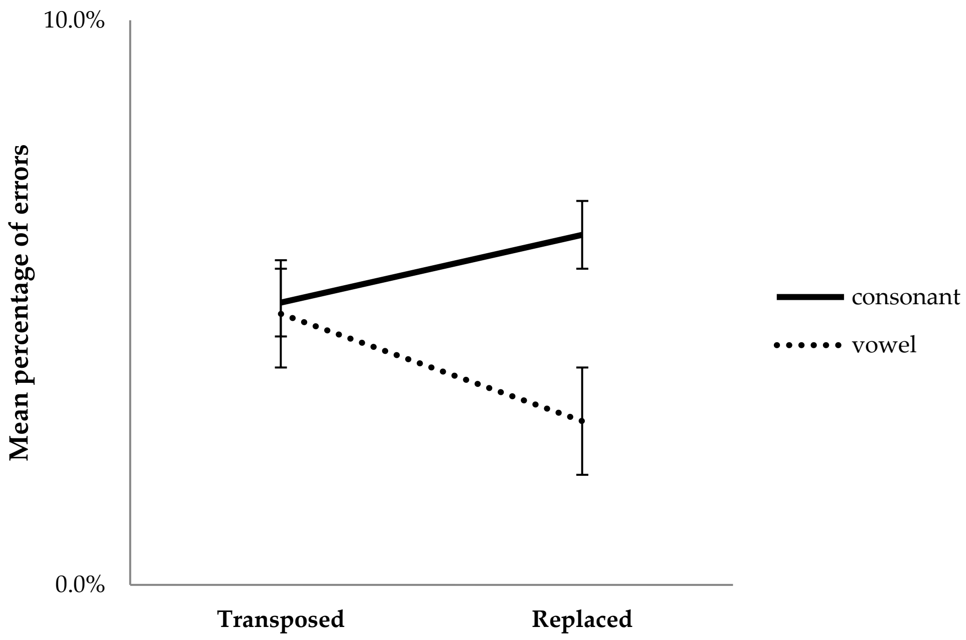

Figure 3.5: Mean percentage of errors for the word prime conditions.

To sum up, error rates for words seemed to follow the same pattern as response times, that is, more errors were observed in the conditions that involved the slower responses.

For comparison, we ran the same analyses with error rates to non-word targets. There was a significant main effect of Phoneme type (F_1 (1, 46) = 32.34, *MSE* = 28.03, p < .001, η_p^2 = .413; F_2 (1, 68) = 11.04, MSE = 65.35, p = .001, η_p^2 = .144), corresponding to an increased error rate for vowel modified non-words than for consonant modified non-words (see Figure 3.6), suggesting that vowel modifications are harder to perceive than consonant changes. Thereafter, none of the effects in the non-word item analysis approached significance in the error data, and so will not be reported any further. There was no effect of Prime type,

(F_1 (1, 46) = .007 MSE = 43.95, p > .05, η_p^2 = .000), and no significant interaction between Phoneme and Prime (F_1 (1, 46) = 1.99, MSE = 34.26, p >.05, η_p^2 = .042).

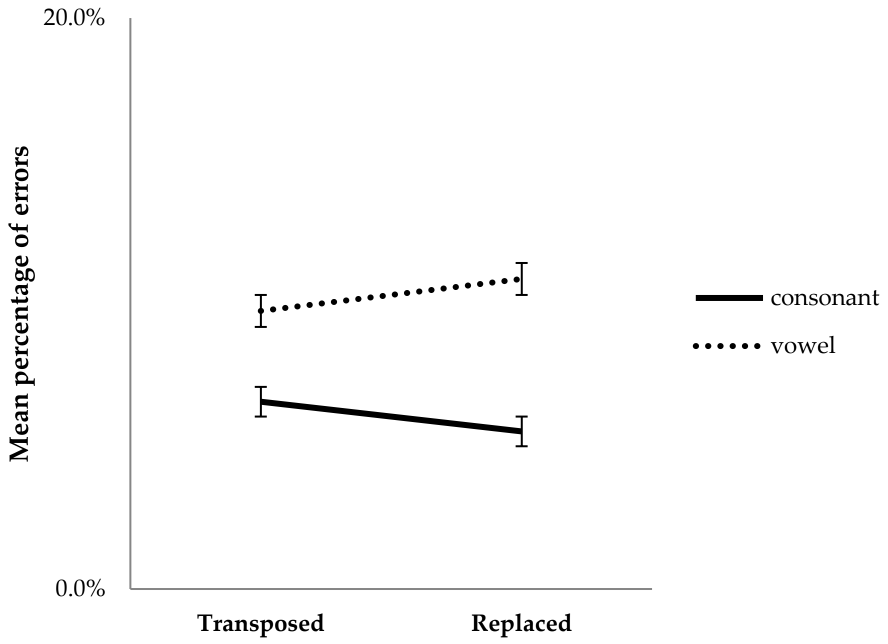

Figure 3.6: Mean percentage of errors for the non-word prime conditions.

3.1.2 Discussion of Experiment 1

To shed some light on the conflictual results obtained in replaced and transposed letter/phoneme paradigms in the visual and auditory modalities (Delle Luche et al., 2014; Lupker et al., 2008; New et al., 2014), we compared directly TL and RL in an auditory experiment. Specifically, a phonological priming experiment using transposed and replaced consonant and vowel primes was used to explore the C&V asymmetry in English adults. In this way, we could also target the use of phonological information, and control for any phonological ambiguity between primes and targets, which might have arisen with Lupker et al. (2008) visual material.

We found an overall advantage of processing consonants over vowels in auditory lexical access: words preceded by primes obtained by modifying vowels and preserving consonants were recognised faster than those preceded by primes modifying consonants and preserving vowels. The same pattern of results was found for accuracy. This clearly departs from the original visual task of Lupker et al. (2008) in which consonant and vowel modified stimuli were processed equally fast. Overall, our findings provide supportive evidence for the consonant bias hypothesis (Nespor et al., 2003), and agree with findings in recent spoken word processing (Delle Luche et al., 2014) and visual priming work (e.g. New et al., 2008).

Interestingly, an interaction between phoneme type and prime type was found, so that consonant replacement impaired lexical access more than consonant transposition, whereas the reverse was found for vowels: replacing vowels impaired word recognition more than transposing them. We shall return to this point in the general discussion.

3.2 Experiment 2: Transposed and Replaced Letters

Following on from our findings in Experiment 1 and given that we had made several modifications to Lupker et al.'s (2008) original stimuli, the aim of Experiment 2 was to re-run Experiment 1 in the visual modality, to show that the results replicate Lupker et al.'s and differ to that of Experiment 1. Here we expected no significant differences between consonant and vowel changes and, as found in Lupker et al., the only faster condition should be the consonant transposition.

Method

Participants

Forty-six monolingual participants, 25 females and 21 males (mean age: 23.0, SD: 6), were tested for the current study. They were recruited with the same criteria as in the previous experiment.

Stimuli

The stimuli and design were the same as used in Experiment 1 (which included the same slight modifications as compared to Lupker et al., 2008). The only difference was that this experiment was adapted to the visual domain.

The word targets in Experiment 2 were 65 English words (as we removed the 15 target words which had not been properly recorded in Experiment 1). The mean length was 6.9 letters (range 5-9) and their mean word frequency per million in the CELEX count was 34.3 (Baayen et al., 1996).

All targets for words and pseudo-words were presented in uppercase letters and were preceded by primes in lowercase that came from one of the four pseudo-word prime conditions. All the stimuli were presented in 18-point bold black text, in Courier New (a monotype font) on a white background screen.

Procedure

The procedure was **identical** to that used in Experiment 1 (e.g., responding with both hands and pressing the same buttons) apart from the stimuli being visually presented. Thus, the procedure is the same as Lupker et al (2008). Response times were measured from target onset until the participant's response. Participants were seated approximately 45 cm in front of a 17" LCD monitor screen. For each trial, a fixation cross was presented in the centre of a screen for 1.5 seconds, followed by a forward mask consisting of a row of six hash marks (######) presented for 500 ms in the centre of the screen. Then, a centred lowercase prime was presented for 47 ms, as in Lupker et al. (2008), which was then replaced by an uppercase target item, which remained on the screen until the response was given (see Figure 3.7 for a visual illustration of the procedure).

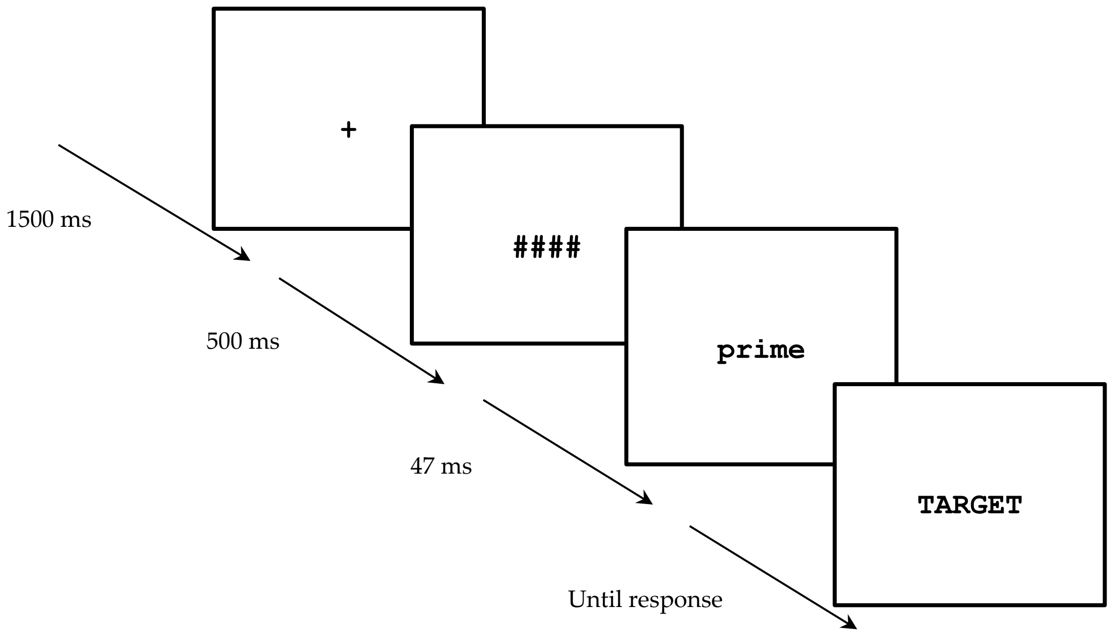

Figure 3.7: An illustration of the procedure used for a typical trial during the lexical decision task.

Data processing

For word data, the response time data to 162 incorrect responses (5.4%) were removed. Response times greater than 1334 ms (cut-off corresponding to 1% of the total trials) including responses below 250 ms were discarded (29 trials). The response times that fell above and under 2.5 SD from the mean for each participant (76 trials) were outliers and excluded from the analyses (2.7%).

Likewise, for non-word data, the response time data to 182 incorrect responses (5.7% of the trials) and response times greater than 1899 ms (cut-off corresponding to 1% of the trials) including responses below 250 ms were excluded from the analyses (29 trials). This was followed by removing all outliers above and less than 2.5 SD from the mean for each participant (83 trials, 2.8% of the trials).

3.2.1 Results

Words (M = 571, SD = 94) were unsurprisingly processed faster than non-word (M = 671, SD = 156; t_1 (45) = -6.98, $p < .001$), and will be analysed separately.

Word Data

Response times were analysed with a 2 (Prime type: transposition, replacement) X 2 (Letter type: consonants, vowels) X 4 (List: list 1, list 2, list 3, list 4) design.

List had a main effect in the item analysis (F_2 (3, 64) = 18.29, $p< .001$, η_p^2 = .462) but not in the participant data (F_1 (3, 42) = 1.34, $p = .27$, η_p^2 = .088). Since it didn't interact with any other factor, it was removed from further analysis.

Table 3.6 depicts reaction times for each Letter type and each Prime type, together with Lupker et al.'s (2008) values in italics. In our experiment, the most important difference between conditions was between the two consonant-modified conditions, corresponding to faster response times to consonant transpositions than to consonant replacements (35 ms); to a lesser extent, vowel transpositions were processed faster than vowel replacements (13 ms).

Table 3. 6: Mean lexical decision times (in ms) and Standard Error (in parentheses) for F_1. Lupker et al. (2008) mean lexical decision times are reported in the right-hand side of the split cell in italics.

Letter Type	**Type of Prime**					
	Transposed		Replaced		Condition effect	
Consonant	554 (*14.3)*	*639*	589 (*14.9*)	*663*	35 (0.6)	*24*
Vowel	564 (14.*3*)	*650*	577 (*16.2)*	*653*	13 (1.9)	*3*

The repeated measures ANOVA showed a main effect of Prime type, (F_1 (1, 45) = 17.09, p < .001, η_p^2 = .275; F_2 (1, 67) = 7.35, p < .01, η_p^2 = .099). This was due to a slower identification of targets after the replaced primes (F_1: M = 583.2, SD = 102.6; F_2: M = 580.4, SD= 54.7) than the transposed primes (F_1: M = 559, SD = 90.8; F_2: M = 563.3, SD= 58.9). The main effect of Letter type was not significant (both p >.05), showing that consonant changes were processed as fast as vowel changes.

As illustrated in Figure 3.8, there was a significant interaction between Letter and Prime (F_1 (1, 45) = 17.09, p = .021, η_p^2 = .113; F_2 (1, 67) = 3.94, p = .048, η_p^2 = .057), due to a larger transposed-letter – replaced-letter difference in the consonant condition (35 ms; t_1 (45) = 6.63, p < .001; t_2 (67) = 3.30, p = .002) than in the vowel condition (13 ms; all p >.05). However, the interaction between the

Prime and Letter in the item analysis with the discarded missing values did not approach significance (F_2 (1, 59) = 2.15, $p > .05$, $\eta_p^2 = .035$).

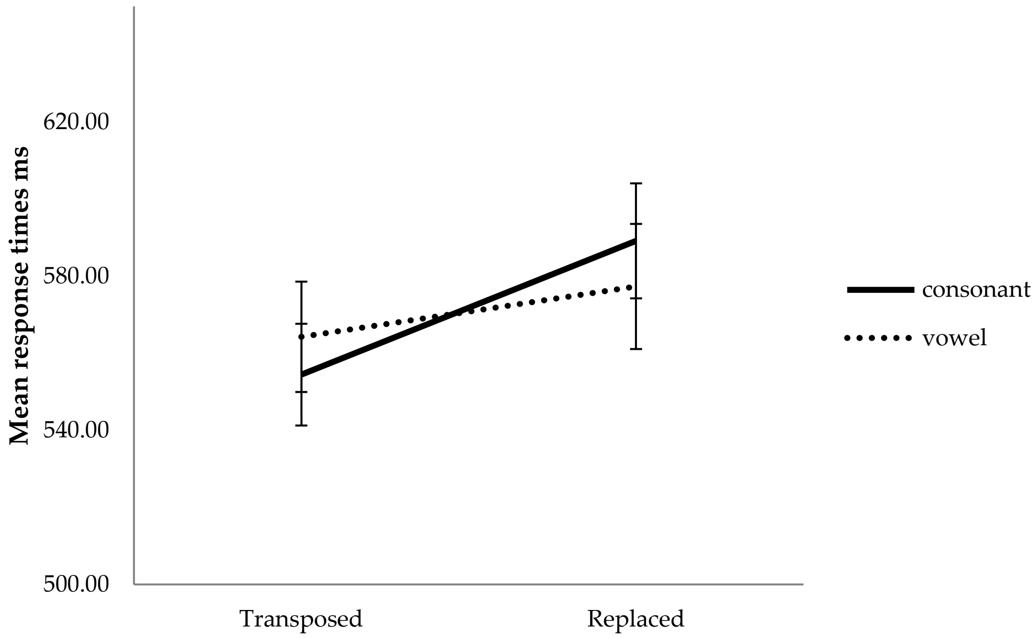

Figure 3.8: Mean response times for the prime conditions. Error bars represent standard error of the mean.

In summary, for consonants, as in Experiment 1 in the auditory modality, transposing letters facilitates word recognition more than replacing them. For vowels, Experiment 2 showed no significant difference between transposing letters and replacing them. This stands in contrast to Experiment 1 in which a significant advantage of replacing vowels over transposing them was found. Importantly, there was no overall advantage for vowel changes over consonant changes, as found in Experiment 1.

Non-Word Data

Again, List had a significant main effect in the item analysis (F_2 (3, 68) = 12.64, p < .001, η_p^2 = .358) but not in the participant analysis. Since list didn't interact with any other factors it wasn't included in further analyses. Letter and Prime did not have any effect nor interacted with one another, as illustrated in Figure 3.9.

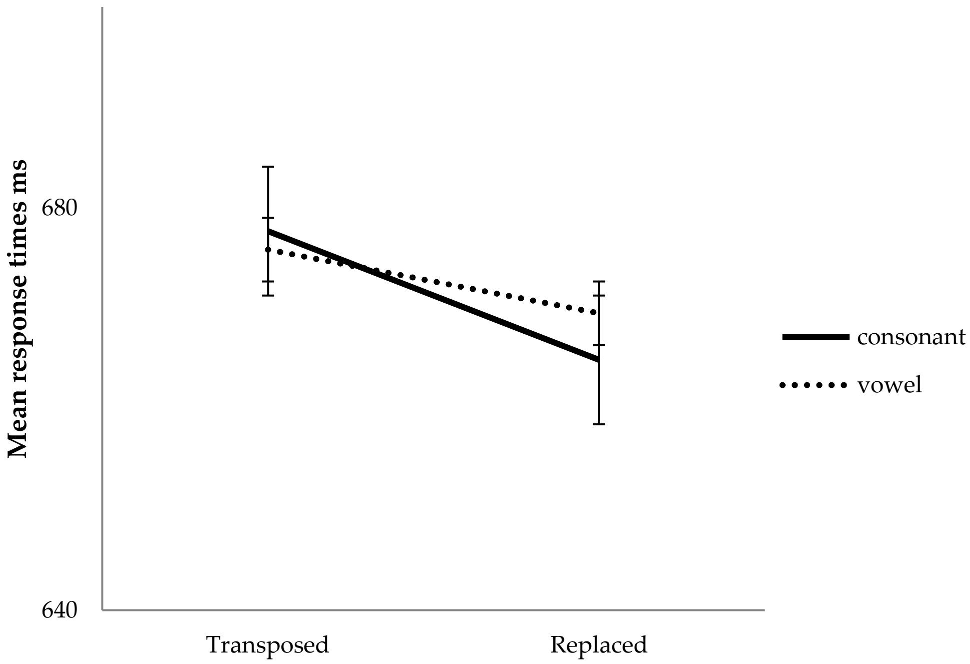

Figure 3.9: Mean response times for the non-words in Experiment 2.

Error Rates

There was no effect of List between groups in the participant data (F_1 (3, 42) = .100, MSE = 52.53, *p* = .959, η_p^2 = .007; F_2 (3, 64) = .087 MSE = 69.21, p = .967, η_p^2 = .004). There was a significant interaction between Letter and List (F_1 (3, 42) =

4.12, $p = .012$, $\eta_p^2 = .228$). However, no main effect of List or interactions were found in the item analysis (all F_2 = >.05), and so these will not be discussed any further.

There was no effect of Prime type (p >.05) but a main effect of Letter type (F_1 (1, 42) = 4.64, MSE = 22.19, $p = .037$, $\eta_p^2 = .100$. As can be seen in Table 3.7, overall consonant changes produced more errors than vowel changes.

Table 3.7: Percentage of errors and Standard Deviations (in parentheses) for word targets. Lupker et al.'s results (2008) are reported on the right-hand side of the split cell for a comparison.

Letter Type			**Type of Prime**	
	Transposed		Replaced	
Consonant	4.6 (*0.6*)	3.3	7.7 (*1.1*)	3.1
Vowel	4.9 (*0.9*)	4.3	4.2 (*0.6*)	2.9

None of the main effects or interactions for error rates in the non-word data approached significance.

3.2.2 Discussion of Experiment 2

Experiment 2 was a direct replication of Lupker et al.'s paradigm (2008), using our stimuli which had been slightly modified from theirs due to the adaptation to the auditory modality in Experiment 1. Overall, our results are very similar to those found by Lupker et al: first, there was no main advantage of consonant-preserved primes over vowel-preserved primes. Second, in the consonant condition, there was a significant transposed-letter advantage (in comparison to the replacement-letter primes). The slight modifications in the stimuli lists that we performed when adapting Lupker et al.'s stimuli to the auditory modality in Experiment 1, and which have been carried forward to Experiment 2, did not modify the global pattern or results.

3.3 Discussion of Adult Experiments (1 and 2)

The aim of this study was to examine the status of consonants and vowels in adult lexical processing. Experiment 1 aimed to compare how replaced and transposed consonant and vowel primes would affect lexical access in English adults whilst listening to spoken words. In direct contrast to Lupker et al. (2008), our results showed a clear-cut consonant-bias priming effect in English words. This replicate and extends the visual consonant bias effect observed in French adults reported by New et al. (2008) and followed the auditory consonant bias effect observed in French and English adults by Delle Luche et al. (2014); it also provides further evidence that the difference in priming observed between consonants and vowel occurs at the phonological level. Moreover, by showing an absence of consonant bias with non-word targets, it suggests that the consonant bias is related to lexical access.

In Experiment 2 we tested priming effects in the visual modality, in a direct replication of Lupker et al., to examine whether our (minor) changes in stimuli would still allow us to observe a pattern similar to theirs. We found here exactly the same results as Lupker et al., namely that there was no main advantage of consonant-preserved primes over vowel-preserved primes, and that there was a significant consonant-transposed letter advantage (in comparison to the consonant-replacement letter primes).

Taken together, it suggests that when we tap into phonology associated with written English words by using auditory stimuli, a different pattern emerges in the two modalities. Spoken word recognition results in a consonant bias as was found in Experiment 1 and in Delle Luche et al. (2014). Overall, in visual word recognition, the C-bias emerges above and beyond the C-TL effect which was only found at the orthographic level. Indeed the consonant bias effect emerges with longer prime presentation (66 ms) but not with shorter ones (33 ms) as reported by New et al. (2014). When the prime duration is at 50 ms (as in New et al., 2008), or 47 ms as in Lupker et al. (2008) or here, no consonant bias is found. This outcome suggests that the nature of the C-bias is phonological, and the nature of the C-TL effect is orthographic (see Chapter 2.2.2).

The second interesting results emerging from Experiment 1 is the significant interaction between prime type and phoneme type: replacing consonants was found to impair lexical access more than transposing them, whereas replacing vowels was less weakening than transposing them. The first component of this interaction is in essence like Lupker et al. (2008) in the visual modality (and in Experiment 2) and confirms that consonant transposed primes lead to a greater lexical activation of the original target than replaced consonant primes. Spoken word recognition models such as TRACE (McClelland & Elman, 1986) or NAM (Luce, 1986; Luce & Pisoni, 1998) seem particularly suitable to account for these findings. Indeed, in TRACE, activation of word candidates that share any similarity with the target words occurs at any moment, and no mismatch inhibition is involved (lateral inhibition between words is the regulating principle).

Therefore, when hearing the prime /adakemy/, the listener would activate the target word ACADEMY at any incoming phoneme, as ACADEMY shares all its phonemes with ADACEMY. In contrast, upon hearing ABANEMY, the target word ACADEMY would still be activated but other competitors sharing the phonemes /b/ and /n/ are also included in the set. NAM differs slightly as it predicts that only those words that differ by no more than one phoneme from the target word will be activated, relying on the idea of recognition based on global similarity. A first reading of this would be that both transposed and replaced consonant primes should activate equally (poorly) the target word as they differ from the target by exactly 2 phonemes. However, NAM's metrics of phoneme-to-phoneme similarity does not consider temporal locations of common phonemes (Magnuson et al., 2007). That would predict more activation for the target word in the case of transposed consonant primes rather than replaced consonant ones.

The other component to the interaction, namely, that more target activation occurs after replaced vowel primes than after transposed vowel primes, was found only in Experiment 1. This could relate to the observation that switching vowels is often used in languages to signal grammatical category changes as in Semitic languages (McCarthy, 1985).

In English, such examples of the role of vowels are scarce but can be found in irregular past-tense verb inflections as in SING, SANG and SUNG. In our study, transposing vowels could signal to the word recognition system that the prime and the target are contrasted at the grammatical level, which would activate further processing as compared to a situation, such as replaced vowel primes, where there would be no other link than global phonological similarity. To examine this possibility, it would be interesting to compare auditory priming for irregular verbs when the vowels are modified. For instance, as demonstrated in the visual modality, we would predict that SENG might prime SING more than STAP versus STOP because there will be memory traces in the lexicon recognising that this irregular verb has some vowel-change inflections, whereas STOP does not (Pastizzo & Feldman, 2002). Thus, the former prime might activate grammatical processing and signal a position of change and category of words, whereas the latter prime would not.

In summary, the main result of this study is that in the auditory modality, changing vowels results in larger target word activation than changing consonants. Consistent with Nespor et al.'s (2003) hypothesis, our results suggest that consonants play a more important role for the identification of word candidates. Importantly, as well as previous priming studies (Delle Luche et al., 2014; New et al., 2008, 2014) this outcome follows the results from word reconstruction and identification tasks (e.g., Cutler et al., 2002; 2000; Sharp et al., 2005; Van Ooijen, 1996), and lexical segmentation tasks (Bonatti et al., 2005). In addition, our result provides further clarity on the findings by Lupker et al. (2008) who showed no main effect of consonant preserved primes over vowel preserved primes.

In conclusion, the C&V dissociation that has emerged in adult lexical processing suggests that the C&V difference is not specific to the phonetic structure of some languages. Similarly, the C-TL effect found in English and Spanish suggest that the C-TL effect in English is not contaminated by the orthography of English. Using an auditory version of the TL paradigm enabled us to directly investigate the phonological nature of the C-bias and the orthographic nature of the C-TL effect in lexical processing. Overall, we can conclude that the nature of C-bias is phonological, and the origins of the C-TL effect is orthographic.

Chapter 4: Introduction to the Processing of Consonants and Vowels in Children

In Chapter 3, we provided strong evidence of a consonantal bias (C-bias) in online lexical processing in English adults. Whilst the adult C-bias appears relatively stable across languages, its developmental origin is less clear. There are three hypotheses regarding the origin of the C-bias. The "initial bias" hypothesis claims that infants process consonants and vowels at the onset of language acquisition, predicting a C-bias would be present at birth (Bonatti et al., 2005). Consequently, this hypothesis predicts no developmental or cross-linguistic difference. In contrast, the "lexical" hypothesis predicts that the C-bias reflects experience with distributional information at the lexical level (Keidel et al., 2007). Thirdly, the "acoustic/phonetic" hypothesis predicts that the C-bias reflects experience with the acoustic-phonetic properties of consonants and vowels in a language (Floccia et al., 2014). Thus, the last two hypotheses predict the bias is learned. Indeed, disentangling age-related predictions is not straightforward as interactions could be observed constrained at a language-specific phonotactic level. The next experimental chapter (Chapter 5) will compare consonants and vowels in English toddlers, to explore the origin of the C-bias.

The main goal for this literature review chapter is to introduce the different positions around the emergence of the C-bias. Therefore, the existing cross-linguistic behavioural research which has explored consonants and vowels in early lexical representations will be presented. In doing so, we will also address positional effects in the emergence of the C-bias.

4.1 C&V in rules and words in toddlers ALL

The CV hypothesis predicts that consonants and vowels serve distinct functional different roles in language (Nespor et al., 2003). First, consonants over vowels are proposed to be more beneficial at the lexical level. Secondly, vowels are more important for processing grammatical information. The focus of the current subsection is to review the literature which has explored the second part of the CV hypothesis, namely, that vowels are more important for aspects of syntax (Hochmann et al., 2011; Pons & Toro, 2010). That said, Hochmann et al (2011) also tested the first part of the CV hypothesis, exploring whether infants are more reliant on consonants than vowels in lexical processing and so this will also be discussed.

Pons and Toro (2010) hypothesised that if consonants and vowels signal different linguistic strres as demonstrated with adults (e.g., Toro, Nespor, Mehler, & Bonatti, 2008)[13], then infants might also show similar constraints. Pons et al. (2010) anticipated that if consonants are more useful than vowels for lexical identification, then infants should behave differently on a structure generalization task (e.g., Marcus, Vijayan, Rao, & Vishton, 1999) when

[13] The adult evidence was discussed in Chapter 2.

implemented over consonants instead of vowels. In an influential study, Marcus et al. (1999) had argued that since 7-month-olds could discriminate between grammatical structures where vowels were arranged according to an ABB rule (/wo/ /fe/ /fe/), or an ABA rule (/wo/ /fe/ /wo/), babies could extract syntactic regularities by learning the rules that generate them. According to the CV hypothesis, this is possible because vowels play a more important role than consonants for encoding aspects of syntax (Nespor et al., 2003). Following, Pons et al. (2010) predicted that infants should fail at this task when the rule is implemented over consonants but succeed when applied to vowels, and they should be sensitive to this distinction before they have a fully developed lexicon.

To test this prediction, Pons et al. (2010) presented Spanish 11-month-olds with a series of non-words with vowels arranged according to an AAB rule in a preferential task. So, the first and second vowel were the same, while the third vowel was different. Similarly, this procedure was repeated for consonants in a second experiment. After a familiarization phase containing CVCVCV nonsense words in which the vowels conformed to the AAB structure, infants were presented with new non-words for testing. The test items were composed of the same phonemes used in familiarisation, but their combination order was different. They found that Spanish infants could discriminate non-words that respected the AAB structure when implemented over the vowels (e.g., /batalo/, /linide/, /noloda/) compared to those that did not conform to the structure (e.g., /bitado/, /lanude/, /nedota/). When presented with the same AAB rule but executed over consonants (e.g., /didola/, /lilune/, /ninube/) they showed no discrimination compared to the trials with a different consonantal structure (e.g., /dutani/, /litedo/, /nelobi/). Following the adult results (Toro, Nespor, et al., 2008), whilst infants used vocalic information to generalize simple structures, they failed to do so with consonantal information. This implies that vowels are more salient for rule-extraction, which is the same type of learning required for processing grammatical information such as syntax. Overall, this sensitivity seems to exist even before infants have a fully developed lexicon, implying an early vocalic bias in syntax processing which supports the CV hypothesis.

However, because Spanish only contains five vowels, Pons and Toro (2010) used the same vowels and consonants in both familiarization and test phases. This led Hochmann et al. (2011) to highlight that this might not actually reflect an infant's ability to generalize the AAB structure, but instead show their ability to learn and memorise repeated vocalic information in the orders. Hochman et al (2011) proposed that to demonstrate generalisation, vowels and consonants not used in familiarisation should be used in the test phase. Italian has two more vowels than Spanish, and so on this basis they tested Italian 12-month-olds by using a Switch task paradigm with consonant and vowel items that were not used in familiarisation. The authors tested the first part of the CV hypothesis, exploring whether infants are more reliant on consonants than vowels in lexical processing. To do so, infants were taught that one word would predict a toy appearing on one side of the screen (e.g., /dudu/), and another word would predict the appearance of a toy on the other side of a screen (e.g., /keke/). Infants were then tested with a new word that was either created with the consonants of the former word and the vowels of the latter (e.g., /dede/) or vice versa (e.g., /kuku/). During the test phase, no toy appeared. The rationale was if infants assign a lexical role to consonants, then they should look for the toy on the side predicted by the first word. Indeed, the results showed an overall accuracy at test which suggested that infants regarded two words that share consonants more similar than two words that share vowels, i.e. /dudu/ is more like /dede/ than /kuku/. In conclusion, 12-month-olds seemed to find consonants more useful when distinguishing amongst words, which suggests that consonants over vowels are more important in lexical processing.

Hochmann et al. (2011) then examined the second part of the CV hypothesis. This experiment was designed to explore if vowels are more beneficial for detecting and generalising repetition structures than consonants when using different exemplars at test (cf. Pons & Toro, 2010). This experiment was very similar to the first experiment, varying only in details that should differentiate a word-learning task from a structure-generalization task. That is, instead of searching for a toy in the location that was predicted by the consonants (vs. vowels) of the ambiguous word, infants searched for the toy in the location predicted by the structure of the word. This simple structural generalisation is said to reflect an aspect of syntax. In this way, six items for familiarisation contained a consonant repetition (e.g. /lula/, /lalo/, /dado/, /dodu/, /fufa/ and

/fofu/) which was followed by a toy appearing on one side of the screen. Another six items contained a vowel repetition (e.g., /dala/, /dolo/, /fodo/, /fudu/, /lafa/ and /lufu/) which was followed by a toy on the other side. They then tested for generalisation, observing if babies searched for the toy when hearing new words that respected the consonant regularity (e.g., /kike/ and /memi/) or the vowel regularity (e.g., /meke/ and /kimi/). It was predicted that generalisation should be implemented over vowels, not consonants. Indeed, they found that Italian 12-month-olds were better at extracting a repetition-based structure over vowels than consonants. They argued that this result could not be explained by memory or statistical dependencies between syllables but could be accounted for by the CV hypothesis of a vowel-based rule mechanism. Above all, vowels over consonants were found to be more beneficial for generating a rule-based structure, so play a more advantageous role for aspects of grammatical processing.

Furthermore, by using the exact same paradigm and stimuli (Hochmann et al., 2011), Hochmann, Benavides-Varela, Fló, Nespor, and Mehler (2017) showed that Italian 6-month-old babies rely more on vowels than consonants when learning two novel words. In contrast to 12-months of age, at 6-months they failed to show an overall preference for searching the toy predicted by the consonantal information over the vowel information. However, the analysis for their first fixations indicated that the initial response was to rely on vowels. Next, they were tested in an exact replication of Hochmann et al's (2011) second experiment. This time, and in contrast to 12-month-olds, no evidence of generalisation emerged in either the consonant or vowel structure. To investigate if infants were hindered in their performance by having to learn two structures, they were then tested in a between-participants design. In this situation, it was found that infants in the vowel-repetition condition could successfully generalise the rule compared to those in the consonant-repetition who failed to generalise the structure.

In summary, Hochmann and colleagues demonstrated that 12-month-olds profit from a partial division of labour between consonants and vowels. That is, distinct speech categories appear to facilitate the development of the lexicon (consonants) and aspects of syntax (vowels) in parallel. In addition, because 6-month-olds showed a vowel bias (V-bias) when distinguishing between words, the authors concluded that a transition from a V-bias to a C-bias occurs during their second semester. When tested exclusively with vowels on the generalisation of structural regularities, 6-month-olds succeeded. However, whether the vocalic preference emerged due to a functional specialisation role of vowels, or because of its greater acoustic saliency, remained unclear.

To conclude, the Spanish and Italian developmental evidence both support the CV hypothesis (Nespor et al., 2003). It was claimed that since vowels carry prosody which marks more abstract constituents, they provide more information about structural relations (Hochmann et al., 2017; Pons & Toro, 2010). In contrast, the C-bias emerges through the distributional and physical properties of consonants which make them preferable in word-learning processes (Hochmann et al., 2011). Thus, these authors argue that processing biases including the C-bias emerge via constraints imposed on their functions at the onset of language acquisition.

4.2 C&V in early word forms

It was established that French 5-month-olds could detect vowel changes that occurred in their own name better than consonant changes (Bouchon et al., 2015). On top of the finding that vowel changes were detected, the result demonstrated that discrimination was predicted by acoustic factors such as spectral distance. However, by using a similar task it was found that whilst English 5-month-olds failed to detect a single phonetic change (consonant and vowel) within their own name, they seemed to find it easier to detect a consonant change with a high intensity contrast (such as a plosive versus a fricative) than a change involving less intensity difference (such as two plosives) (Delle Luche et al., 2017). This demonstrates that at an age still within the period of so-called universal perception, French and British English-learning 5-month-olds rely on different sets of acoustic cues in early speech perception for forename recognition. Together with the Italian data (Hochmann et al., 2017), the French outcome demonstrates an early V-bias in early word forms, indicating that the C-bias must be learned (Bouchon et al., 2015). Overall, it was suggested that the specificity of word representations and the lexical processing biases reported in older infants might result from language-specific acoustic biases, combined with a lexically-driven learning process.

A C-bias in lexical processing for slightly older infants was reported for French 11-month-olds (Poltrock & Nazzi, 2015). Using an auditory word recognition head-turn task (HTP) they first showed that infants preferred to listen to a familiar word e.g., /bird/ which in French is /wazo/ ("oiseau") compared to a non-word /walØ/. Then, in a conflict situation where babies' preference for a consonant mispronunciation (MP) as in /wavo/ versus a vowel MP as in /wazu/, infants showed a preference for the vowel MPs. This implied that a consonant MP disrupts recognition of a familiar word more than a vowel MP. Hence, toddlers find consonants more useful than vowels in word recognition. Overall, Poltrock et al. concluded that their evidence is compatible with the idea that the functional roles of C&Vs emerge during the second year of life due to differences at the acoustic/phonetic levels, and possibly at the lexical or pre-lexical levels (e.g., Andics, 2006) which would lead to C-advantage for lexical processing.

Nishibayashi and Nazzi (2016) conducted a series of experiments in a recent French study with 8-month-olds which might lend some support to the pre-lexical and acoustic/phonetic accounts. First, the authors explored the impact of C&V MPs on recognising segmented word forms by using CV monosyllabic target non-words. Critically, these non-words were presented to the babies in the test phase as mispronounced as compared to the familiarisation phase. The words were embedded in a passage at the beginning or towards the end of sentences which were used in the familiarisation phase. The syllables preceding and following the target word were always different so that no syllabic sequence was repeated, thus preventing infants from computing transitional probabilities based on a statistical regularity of the syllable pattern. In this way, babies were presented with either a correct pronunciation (e.g., /ti/) versus a vowel MP (e.g., /te/), or a correct pronunciation (e.g., /py/) versus a consonant MP (e.g., /by/). Here, infants oriented equally to the targets and to both C and V MP conditions, so no evidence in favour of the C-bias emerged. Because control words, which are words not used in the familiarisation phase, might provide a more sensitive measure (Swingley, 2005), infants were then tested on MPs vs. control words. This time, a segmentation effect emerged in the vowel condition whereby babies looked longer to the vowel MPs over the control words. In contrast, babies oriented equally to the consonant MPs compared to the control words. This suggested that babies considered the consonants MPs as different to the targets, which implied that 8-month-olds have a C-bias in recognising word forms.

However, a possible confound was that the consonant mispronunciation always came before the vowel mispronunciations within the CV words, so to address the possibility of a positional effect, a third experiment used a conflict situation (Nishibayashi & Nazzi, 2016). This time, 8-month-olds were tested with real words in one of two conditions where the consonant MP occurred either in the onset (CV) or coda (CVC) position, and vowel MPs occurred in coda or medial positions. The idea was that if the C-bias extends to the coda position then the same pattern should be found in both the CV and CVC condition. The outcome confirmed that the C-bias was not dependent on the position. The lack of a positional effect mirrored previous word-learning observations at different ages and with different tasks in French infants (Nazzi & Bertoncini, 2009; Poltrock & Nazzi, 2015).

Nevertheless, Von Holzen, Nishibayashi and Nazzi (2018) recently published electroencephalographic (EEG) data from an event-related brain potential (ERP) study, indicating that in some tasks consonant onsets might be represented with more detail than codas. Thus, the developmental case of positional effects (onsets vs. codas) in word form segmentation for the emerging C-bias in French requires further testing. Overall, the data from Nishibayashi et al. (2016) lends some support to a possible role of pre-lexical and acoustic-phonetic in the emergence of a C-bias, both of which might not be mutually exclusive (Nazzi, Poltrock, & Von Holzen, 2016).

4.3 C&V in early word-learning tasks

In French, Nazzi (2005) made the first demonstration that French infants weigh consonants more than vowels whilst learning new words. Following the data showing that French adults track transitional probabilities (TPs) amongst consonants but not vowels (Bonatti et al., 2005), Nazzi (2005) investigated the processing of consonants and vowels in the lexicon with French toddlers. Using an adapted version of the interactive name-based categorization task (Nazzi & Gopnik, 2001), 20-month-olds were introduced to novel pairs of objects labelled with a different non-word, e.g., /duk/ and /guk/. Next, they were given a different new object which is also labelled with either /duk/ or /guk/ and were asked to choose from the other two objects, the one that it matches. Thus, the decision was based solely on whether it shared the consonants or the vowel of the target non-word. They found that French infants successfully learned new words based on initial and non-initial consonantal contrasts, but repeatedly failed when asked to do the same task involving vowel changes (e.g., /duk/ versus /dɔk/). Overall, this study revealed that infants were able to consider minimal consonantal differences in either word-initial or embedded in accented syllables. In contrast, their performance on three vocalic contrasts were at chance level. Hence, Nazzi (2005) provided the first piece of developmental evidence for a greater reliance on consonants over vowels at the lexical level.

Furthermore, these results were later extended to include continuous consonants. Continuous consonants are types of sounds in which air flows freely through the vocal tract which is never obstructed e.g., liquid contrasts such as /rize/ and /lize/ (Nazzi & New, 2007), and initial voicing contrasts and consonant coda position in CVC words at 24-months (Nazzi & Bertoncini, 2009), extending earlier findings to include other consonantal contrasts and positions (coda). This result was also replicated with 16-month-olds using a simplified version of the task (Havy & Nazzi, 2009). Taking together, French-learning infants have revealed a C-bias in word-learning at 16-, 20-, and 24-months of age which emerges in the syllable-onset positions of both mono- and disyllabic words, and coda positions in CVC words.

In summary, the French data reveal a robust C-bias in lexical representations which emerges from as young as 8-months of age (Nishibayashi & Nazzi, 2016). However, what is interesting is that the English data reveal a different pattern of results.

Whilst a cross-linguistic comparison between French and English show that both English and French infants exhibit a C-bias at 30-months of age (Nazzi et al., 2009), English infants at 16- to 23-months do not display a C-bias (Floccia et al., 2014). In a direct replication of Nazzi et al (2009) where positional effects were controlled for, Floccia et al. (2014) used CVC non-words where the first consonant (C_1), the medial vowel, or the coda consonant (C_2) were manipulated (e.g., /dib/ vs. /dɛb/ vs. /gib/). In a within-participant design, half of the consonant changes happened on either the initial consonant, or coda-consonant. Thus, half of the trials consisted of pairs such as /dib/ vs. /dɛb/ (vowel change), or /dib/ vs. /gib/ (initial consonant change). In contrast, the other half required a decision based on comparing /dib/ vs. /dɛb/ (vowel change), or /bɒp/ vs. /bɒt/ (final consonant change). The consonant contrasts were chosen based on a single place of articulation change, and the vocalic contrasts involved either height or roundness. Overall, the results showed that English-speaking infants did not show a greater reliance on consonants than vowels when learning new words. However, when consonant contrasts differed in the coda positions such as, /bɒt/ and / bʌp/, and the test word was /bɒp/, infants paired the two objects that shared the consonant information /bɒp/ and /bʌp/ significantly more than when they shared the vocalic information /bɒp/ and /bɒt/. This result indicated that the coda contrast is more salient than the consonant onset contrast.

Floccia et al. (2014) proposed possible explanations for the C_1/C_2 asymmetry. First, it was suggested that it might due to a recency effect (e.g., Burgess & Hitch, 2006) where the last segment is processed better than the preceding one. Yet this explanation would suggest a (medial-) vowel bias over onset consonants and the vowels were found to be no different to the C_1 contrasts. Another plausible recency effect is that the effect could be restricted to the final segment which happened to always be a consonant in the CVC words. But then, an equal sensitivity to onsets and codas in familiar word recognition (Swingley, 2009a), in familiar word mispronunciations (Swingley, 2005) and in interactive word-learning all suggest an equal sensitivity to onset and coda (Nazzi et al., 2009). That said, this interpretation does not account for language-specific differences given the lack of a positional effect in the French word-learning data which revealed a sensitivity in both consonantal onset and coda segments (Nazzi et al., 2009). On this basis, an explanation based on a recency effect was ruled out.

Alternatively, a possible interaction with rhyme sensitivity was considered. Specifically, with C_2 triplets such as /bɒt/ and / bʌp/, and the test word /bɒp/, there is a different rhyme in the labels. In the case of C_1 triplets the target /dib/ shares the rhyme with only the vowel-sharing label /gib/ and not the consonant-sharing label /dɛb/. Toddlers would pair /dib/ with /dɛb/ if consonantal information is processed better than vocalic information. However, sensitivity to rhyme overlap (e.g., Treiman & Zukowski, 1996) would lead them to pair /gib/ with /dib/. Against this backdrop, it is plausible that the opposite trends of rhyme and C-bias could nullify any preference for consonants or vowel pairing.

Lastly, another possible explanation for the C_1/C_2 asymmetry is that English infants do attend to consonants more than vowels in lexical processing but only in word-final position. This explanation is possible when considering children's evidence with vocal production which show that whilst English infants increase their production of coda consonants, their French peers drop theirs (Vihman & Boysson-Bardies, 1994). This could explain why a potential positional effect was found in English and not found in French (Nazzi et al., 2009). To explore the C_1 and C_2 asymmetry, a second experiment used a simplified version of the word-learning task which uses two labels rather than three (Havy & Nazzi, 2009). In this way, a decision cannot be based on labels that share a rhyme or an initial consonant. It was predicted that if the outcome from the first experiment was driven by a combination of a rhyme effect and a C-bias, then no difference between C_1 and C_2 should emerge. Following the procedure from the fist experiment, it was found that the final consonant was not processed better than the vowel, which is different to what was found in French (Havy et al., 2009).

To determine if the C-bias repeatedly found in French lexical processing (e.g., Nazzi et al., 2009) might be particular to the French linguistic input, Floccia et al (2014) explored if language-specific cues, based on the rhythmic properties between French and English, might have influenced behaviour. French is a syllable-timed language and so contains clear syllabic boundaries. In contrast, English is a stress-timed language that contains unclear syllabic boundaries which contribute towards a phenomenon referred to as ambisyllabicity (e.g., Goslin & Frauenfelder, 2001). On that basis, English infants were presented with the same French stimuli as a final test in the Floccia et al., (2014) study and once again the same English-specific result was obtained, suggesting that the C-bias found in French is shaped by experience with the acoustic and phonological properties of their native language.

Overall, the data shows that English children between 16- and 23-months of age do not pay more attention to consonants compared to vowels in word-learning tasks. This contrasts with their French peers who show a robust C-bias in word-learning regardless of its position in a word. Thus, up to now, the trajectory of the C-bias appears to be dependent on the language environment.

Finally, in Danish, a recent word-learning study has shown a reversed V-bias in early language processing. Using an object manipulation task (e.g., Havy & Nazzi, 2009), Højen and Nazzi (2016) reported that whilst 20-month-old Danish infants could learn phonetically similar pairs of words that contrast by a vowel (e.g., /dyl/ - /dul/), they were unable to do so when they contrast by a consonant (e.g., /fan/ - /san/). This bias to the advantage of vowels is the first time an early vocalic lexical processing bias has been shown, which goes against

Nespor et al.'s (2003) claim that the task of distinguishing lexical items rests more on consonants than on vowels. Danish is a language which contains 32 vowels and 20 consonants, making it one of the most highly vocalic languages. In addition to its larger vowel inventory, Danish phonology has distinctive features including extensive lenition (weakening) of consonantal sounds. As a result, Højen et al. (2016) argued that speech processing biases, whether vocalic in the case of Danish or consonantal in French (Nazzi et al., 2009), must arise from either the phonological or lexical properties of the native language (Floccia et al., 2014). Overall, it was argued that the reversed bias obtained in word-learning between French and Danish 20-month-olds, implies that speech processing biases must be developed over the course of language exposure, and so is not orchestrated by an innate or language-general device.

In summary, by using adapted versions of an interactive word-learning tasks the French, English and Danish evidence reveal differences of when and even if the C-bias emerges. So far, it appears that the early emergence of a C-bias is dependent on the properties of a given language. In French, the lack of a positional effect confirms the robust nature of the C-bias in lexical processing. As well, the Italian data which has used a different paradigm shows that the C-bias in lexical processing emerges early on in development. On the contrary, English infants have not yet revealed a C-bias until 30-months of age, and prior to this age, a positional effect suggests that if anything, toddlers pay more attention to consonants than vowels in word-final position. Danish infants have shown a reverse V-bias at 20-months. As a result, processing biases whilst learning words including the emergence (or not) of the C-bias, appear to be age-related and language-specific. As we will review below, Mani and Plunkett (2007) also explored if the emergence of the C-bias in English toddlers is depended on the nature of the task (name-based categorisation vs. mispronunciation) and/or the status of the words (new vs. familiar).

4.4 C&V in early preferential looking tasks

Following the evidence found by Nazzi (2005), Mani and Plunkett (2007) predicted that infants might be more sensitive to consonant MPs than vowel MPs of familiar words. Using a standard intermodal preferential looking task (IPL) with familiar distractors, they tested English 15-, 17-, and 24-month-olds in a mispronunciation task. After demonstrating that medial vowels play a prominent role in CVC word recognition, they next tested if consonants constrain lexical identify more than vowels. In this experiment, infants heard four correct familiar word pronunciations and two consonant and two vowel mispronunciations. The mispronunciations were created by changing one dimension of the consonant (place, voicing) or vowel (height, backness). The familiar target words e.g., /bus/, and the distractor images began with the same consonants e.g., a bike. Thus, infants saw a visual image of a bus and a bike together with either the correct pronunciation /bus/ or a consonant mispronunciation /pus/ or vowel mispronunciation /bas/. Infants were found to be sensitive to both vowel and consonant mispronunciations. In sum, this was the first piece of evidence suggesting that vowel and consonant identity constrains lexical access equally in familiar word recognition. In conclusion, Mani and Plunkett (2007) argued that the difference between the current English data and French (Nazzi, 2005) is not attributed to the novelty of the stimuli but could reside in cross-linguistic differences in the vowel systems of the two languages. Overall, English infants do not profit from consonantal information over vowels when recognising familiar words, at least unequivocally from 18-month-olds. Thus, this finding goes against the predictions made by the CV hypothesis (Nespor et al., 2003).

Mani and Plunkett (2010) later compared consonant and vowel sensitivity in English 12-month-olds. Using the standard IPL task, they also found no advantage for consonant mispronunciations over vowel mispronunciations with infants as young as 12-months. Whilst this outcome reflects those found with slightly older children (Mani et al., 2007), some important differences emerged. One finding was that a sensitivity to vowel-MPs appeared to improve only with increased vocabulary size, whilst consonant-MPs appeared unaffected by increasing vocabulary size. This suggests a differential impact of language experience on sensitivity to vowel mispronunciations and consonant mispronunciations. Subsequently, Mani and Plunkett (2010) proposed that the change from fine-grained acoustic-phonetic representations to broader phonemic representation of vowels may happen later than for consonants. In all, the main outcome is that British English 12-month-olds are equally sensitive to vowel and consonant mispronunciations, suggesting that both play an important role in constraining infant word recognition. Overall, the English evidence using IPL tasks suggests that infants from 12-months of age can recognise some familiar words that differ in a single vowel, and that there is no significant difference in their performance between consonants and vowels. Thus, consonants and vowels appear to constrain lexical access similarly.

Another possibility to explain the English and French discrepancy between Mani and Plunkett's (2007, 2010) studies and Nazzi et al.'s (2009) might not be task-related, due to the status of the word, or to the age range. Instead, it could be due to the position of the consonant contrast such as the coda position (Floccia et al., 2014). Thus, in order to firmly conclude that English infants do not exhibit a C-bias until 30-months of age (Nazzi et al., 2009), consonant contrasts in coda position need to be further explored, given that Mani et al (2007, 2010) only tested consonant mispronunciations in word onset position.

4.5 Summary

Overall, the cross-linguistic data reveal that in early stages of language acquisition, Italian and French infants encode vocalic information significantly better than consonants during the first semester of life (Bouchon et al., 2015; Hochmann et al., 2011). For these two languages the transition from a V-bias to a C-bias whilst processing words and word forms seem to emerge during the second semester of life (Hochmann et al., 2011; Nishibayashi & Nazzi, 2016). However, even though this shift from vowel to consonant preference in word-learning appears to be replicable in syllable-timed languages, it varies in languages with different phonetic properties and phonological patterns such as English (e.g., Floccia et al., 2014) and Danish (Højen & Nazzi, 2016). In Danish, an opposite vocalic bias (V-bias) emerges at the onset of lexical acquisition (Højen & Nazzi, 2016), and in English the emergence of a C-bias is not observed until 30-months (Nazzi et al., 2014).

Firstly, whilst English IPL studies show similar levels of sensitivity between consonants and vowels at 12-, 18-, and 24-month-olds (Mani & Plunkett, 2007; Mani & Plunkett, 2010), studies using adaptations of the word-learning task show what whilst a C-bias emerges at 30-months (Nazzi et al., 2009), it is not yet present at 16- and 20-months of age (Floccia et al., 2014). However, one exception was when consonants differed on the final segment of CVC non-words where it appeared that the consonantal coda was more salient (Floccia et al., 2014). In contrast, when the items differed on their initial consonants, infants showed no preference. Furthermore, the previous English IPL studies only explored vowels with consonants in onset positions of CVC familiar words.

Therefore, the following chapter aims at clarifying the English data by testing infants in a familiar word recognition task involving vowel MPs and consonant MPs in onset and coda positions. Following the evidence that coda MPs are as well specified as onset MPs in 21-month-olds (e.g., Swingley, 2009), we tested English 21-month-olds attention to consonants and vowels in lexical recognition using the standard IPL paradigm whilst controlling for potential recency effects (e.g., Floccia et al., 2014). Until now, the recency of the segment (first or second on the word) and its nature (consonant or vowel) had always been confounded. In Experiment 3a, we neutralised this potential effect by testing onset consonants and coda consonants, as well as medial vowels.

Chapter 5: Investigating the processing of consonants and vowels in English-learning toddlers

To address the ambiguity with the current English consonant and vowel findings in lexical processing, the following study will test English 21-month-olds with a standard IPL procedure using CVC familiar words and familiar distractors. The IPL paradigm provides a direct and precise measure of sensitivity to an infant's online word comprehension (Delle Luche et al., 2015). Typically, word recognition is indexed by infants looking significantly longer at the target image in the post-naming phase compared to the pre-naming phase (e.g., Golinkoff et al, 2013). This experiment tests whether infants are sensitive to mispronunciations created by manipulating the phonemic class (consonant/vowel) and the consonant location (onset/coda) of familiar CVC words. To date, a potential recency effect has never been controlled for in English, and so has always been confounded with phonemic class. Using a standard IPL task (e.g., (Mani & Plunkett, 2007; Mani & Plunkett, 2010) infants were presented with two pictures of familiar objects such as a picture of a ball and a picture of a car along with a correct or an incorrect pronunciation of the monosyllabic target word.

In line with Mani and Plunkett (2007, 2010), we predict that infants will be sensitive to mispronunciations, meaning that they should look longer at the target following the correct pronunciations. Likewise, in line with Mani et al., (2007) we expect that toddlers will display an equal sensitivity to vowel-MPs compared to the consonant onset-MPs. However, children's behaviour with the vowel-MPs compared with the coda-MPs contrast, can go in one of three ways. If English children behave like their French peers (e.g., Nazzi, 2005), and if children do find the consonantal coda contrast more salient than the onset (Floccia et al., 2014) then we might expect toddlers to demonstrate an asymmetrical sensitivity to the onset-MP compared with the coda-MP, revealing a positional effect in the emergence of a C-bias in English toddlers. If on the other hand, children do not find consonants more beneficial than vowels in familiar word recognition (e.g., Mani & Plunkett, 2007), then the consonantal coda contrast will not be more salient than the onset, providing further evidence that both consonants and vowels constrain lexical recognition equally at this age, regardless of their position. So, if infants look longer at familiar objects when presented with the correct pronunciations, with no asymmetry between onset consonant, vowel and coda consonant mispronunciation detection, it would strengthen the current finding (Nazzi et al., 2009) that English-learning children do not show a C-bias in word processing prior to 30-months of age. Finally, if consonant and vowel are processed similarly, but if word recognition is sensitive to positional effect as speech unfolds, we would expect to see onset-MP generating more word identification disruption than vowel-MP, and then by coda-MP.

5.1 Experiment 3a: Onset Consonant, Vowel and Coda Consonant Mispronunciation Detection with Familiar Distractors

Method

Participants

Twenty-four healthy monolingual English-learning toddlers aged 21-months (M = 20.5, including 11 boys) participated in this study. The data of 21 additional infants were rejected, for either being inattentive (5), such as looking behind them or at their feet, or non-completion (4), experimenter error (4), and technology failures (8). All infants had no known hearing or visual problems, no reported developmental delays and were no more than 6 weeks premature which is the standard procedure at Plymouth Babylab. They were recruited via the Plymouth Babylab database.

Stimuli

The stimuli consisted of 48 monosyllabic CVC familiar nouns as understood by children at this age (see Table 5.1) selected from the Oxford Communicative Development Inventory (OCDI; Hamilton, Plunkett, & Schafer, 2000). These words created the 24 targets (mean understanding score = 89.7%) and unrelated but familiar 24 distractors (89.1%). Each infant was presented with all 24 target/distractor pairs once, with one familiar image acting as the target for all children, e.g. target /cat/ and distractor /plane/. All images consisted of colour photographs, controlled for size, and appeared on a white background on a 52”

TV screen. They measured 36cm diagonally from corner to corner and were presented 31cm apart. They were deemed good exemplars by the authors and independent observers who were Babylab undergraduate researchers. Out of the 24 test trials, there were 6 of each of the 4 pronunciation types: correctly pronounced, mispronounced on the onset-consonant (C1), or the medial-vowel (V), or the coda-consonant (C2), by a single feature where possible. A full list of the word pairings and mispronunciations, including the visual familiar distractors can be seen in Table 5.1. A female native British English speaker produced the stimuli in an infant-directed style. All auditory stimuli were presented in the carrier phrase *"Look! Target word".* The decision to use this carrier phase was based on following previous 'word recognition/learning) IPL studies (e.g., Mani & Plunkett, 2008). Using simple grammatical sentences in the style of infant directed speech is acceptable for the purpose of 'word' studies. Of course, it would be incorrect to adopt this style of carrier phase if we were exploring grammar.

Table 5. 1: List of stimuli with IPA transcriptions for the targets used in all trials. Percentages are the number of 21-month-olds who know the target and distractor words based on the OCDI norms (Hamilton et al., 2000).

Word	Target	%	Onset change	Vowel change	Coda change	Distractor	%
Ball	bɔ:l	100	gɔ:ll	ba:l	bɔ:n	Chicken	89
Bath	bɑ:θ	94	dɑ:θ	bɔ:θ	bɑ:s	Cow	94
Bed	bɛd	94	pɛd	bʌd	bɛg	Spoon	94
Bib	bɪb	78	dɪb	bɛb	bɪp	Carrot	78
Bin	bɪn	78	dɪn	bɛn	bɪm	Clock	82
Boat	bəʊt	89	pəʊt	baʊt	bəʊk	Chips	89
Book	bʊk	100	pʊk	bɪk	bʊt	Nose	100
Bus	bʌs	94	pʌs	bæs	bʌθ	Door	94
Cat	kæt	100	gæt	kɛt	Kæd	Plane	100
Coat	kəʊt	89	təʊt	kaʊt	kəʊp	Fork	89
Cot	kɒt	94	tɒt	kɔ:t	kɒp	Bike	94
Cup	cʌp	83	tʌp	kɛp	kʌb	Flower	83
Dog	dɒg	100	bɒg	dʊg	dɒd	Cake	89
Doll	dɒl	83	gɒl	dɔ:l	dɒn	Finger	83
Duck	dʌk	94	gʌk	dæk	dʌt	Train	83
Fish	fɪʃ	89	vɪʃ	fɛʃ	fɪʒ	Bottle	83
Foot	fʊt	83	θʊt	fɪt	fʊp	House	89
Hat	hæt	78	ʃæt	hɛt	hæp	Slide	78
Keys	ki:z	78	ti:z	ku:z	ki:v	Plate	88
Leg	lɛg	89	nɛg	lɪg	lɛk	Bunny	89
Pen	pɛn	83	bɛn	pæn	pɛm	Horse	100
Pig	pɪg	100	tɪg	pɛg	pɪd	Car	100
Sheep	ʃi:p	89	ʒi:p	ʃu:p	ʃi:b	Button	83
Sock	sɒk	94	zɒk	sɔ:k	sɒt	Bear	89

Procedure

Parents completed the OCDI (Hamilton et al., 2000) prior to the visit. After ethical consent was obtained children were tested individually accompanied by their caregiver in a quiet room. During the experiment, all infants sat in a highchair approximately 80 cm away from the television screen. Two cameras positioned directly above the visual stimuli recorded infants' eye-movements. Software recorded time-locked images of the infants looking behaviour for the duration of the task. Auditory stimuli were presented via a central speaker. The experiment was created, presented, coded and analysed with the Lincoln Infant Lab software package (Meints & Woodford, 2008). Following two training trials which were always correctly named (hand-chair, bird-mouse) infants were each presented with 24 test trials.

In each trial, infants saw images of two familiar objects side-by-side on a screen for 5 s. The target object was named in the carrier phase, i.e. "look /cat/", with the onset of the target word occurring at 2500ms, splitting the trial intro pre- and post-naming phases. A central fixation smiley-face emoji was presented between trials to centralise the infant's attention. The targets were presented equally often to the left and right, and correct and incorrect pronunciations equally often to the left and right. The order of trial presentations was counterbalanced and randomised so that no more than two correct pronunciations occurred consecutively, and mispronunciations from the same condition did not occur consecutively. The stimuli lists were created so that each child saw only one pair out of the 4 conditions, for any target/distractor pair.

Scoring

Videos were scored to determine the infants' gaze direction and fixations on a frame-by-frame basis (every 40ms). The coded frames were used to calculate the amount of time toddlers spent looking at the target and distractor in each of the pre-and post-naming phases for each trial. As in previous research, looking times that occurred between 367 ms and 2000 ms after the onset of the target word were analysed (Swingley, 2009a). As well, the inclusion criteria per trial was that at least one image had to be fixated on during the pre- and post-phase, and children must also know the familiar distractor. Out of 576 trials, 101 were excluded on that basis which left a total of 475 for the analysis. Thus, 82% of all trials were retained.

5.1.1 Results

To obtain a proportion of target looking time (PTL), we calculated the amount of time infants spent looking at the target (*T*) divided by the total amount of looking at both target and distractor (*T*/*T*+*D*), in each phase. A significant increase in PTL in the post-naming phase compared to the pre-naming phase is taken as evidence that the infant has recognised the word and knows the relationship between the target label and target image, corresponding to a naming effect (Mani & Plunkett, 2010; Swingley & Aslin, 2000).

A repeated measure ANOVA on PTL with Naming (pre and post) and Pronunciation (correct and incorrect) as within-subject factors revealed a main effect of Naming, F (1, 23) = 7.96, p = .010, η_p^2 = .257. This naming effect indicates that infants show an increase in looking times in the post-naming phase (mean .56) compared to the pre-naming phase (mean .50). However, no overall main effect of Pronunciation type, F (1, 23) = 1.65, p = .21, η_p^2 = .067, and no interaction between Naming and Pronunciation were found, F (1, 23) = .298, p = .590, η_p^2 = .013, suggesting that children looked longer post-naming regardless of the target's pronunciation.

Given the past literature on this topic, planned comparisons were conducted to compare target looking between the pre-naming and post-naming phases between the CP and MP (all conditions collapsed), and then for each MP condition. It showed that looks to the target significantly increased from the pre-naming phase to the post-naming phase following the correct pronunciations, t (23) = -2.28, p = .03, d= - 0.46, but not for the incorrect pronunciations, t (23) = -1.67, p = .11, d= - 0.33 (see Table 5.2).

Table 5. 2: Experiment 3: Means (Standard Deviations) for the PTL measures for CP and MPs collapsed together.

Pronunciation	Naming Pre	Naming Post	Effect
CP	.52 (.13*)*	.58 (.16)	.06 (-*0.03*)
MP	.49 (*0.05*)	.53 (*0.09*)	.04 (-0.04)

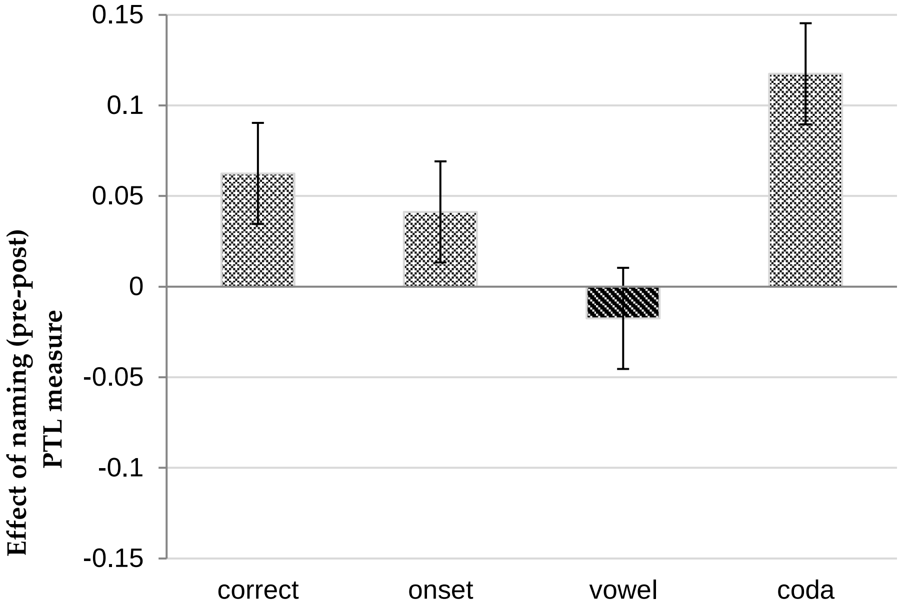

Figure 5.1: Effect of naming (pre-post) in CP, MP-Onset, MP-Vowel, and MP-Coda trials using the PTL measure. Error bars indicate the standard error.

A visual inspection of Figure 5.1 suggests a difference in toddlers' preference for the target between pre- and the post-naming phase when the target label was mispronounced on the coda position. Table 5.3 presents the PTL measures for all types of pronunciations in both the pre-naming and post-naming phases. Post hoc paired-samples *t*-tests confirmed a significant increase in target looking from the pre-post naming phase for coda-MPs, t (23) = -3.34, p =.003, d= -0.68. The only other significant increase that emerged between the pre- and post-naming was for the correct pronunciation, t (23) = -2.27, p =.03, d= - 0.45. However, post hoc paired-samples *t*-tests on the pre and post (Naming) PTL demonstrated that none of the MPs compared to the CP were significantly different (all > .05, see Figure 5.1).

Table 5. 3: Experiment 3: Means (SD) for the PTL measures for CP and MPs separated.

Phase	CP	Onset-MP	Vowel-MP	Coda-MP
Pre-naming	.522 (.13)	.496 (.16)	.530 (.14)	.429 (.14)
Post-naming	.585 (.16)	.538 (.16)	.513 (.17)	.546 (.12)

5.1.2 Discussion of Experiment 3a

The standard IPL paradigm was used to test for infant's attention to mispronunciations of consonants and vowels of familiar CVC words, and a potential consonantal position effect. We found an overall effect of Naming (the magnitude of change from pre-naming to post-naming), indicating target recognition, but no main effect of Mispronunciation or interaction. Unusually, this finding does not correspond to the classic pronunciation effect observed in MP studies (Mani & Plunkett, 2007). Although previous literature shows that the standard IPL procedure provides a valid and sensitive method, research has shown that the standard procedure is not always the most efficient option to study phonological sensitivity in word recognition (K. S. White et al., 2005).

We decided to modify slightly the paradigm for an adaptation of the mispronunciation task which might be able to provide more sensitivity (K. S. White & Morgan, 2008). The standard IPL uses familiar distractors and the adapted version uses novel distractors. That is, objects that the child would not have a name for yet, i.e., a garlic presser. This version allows to measure a response to the distance between correct and incorrect mispronunciations with more sensitivity. Subsequently, MP effects should vary as a function of the similarity between distractor and target (Aslin & Fiser, 2005). If the mispronunciation fits better to the target label than the distractor label, the infant might continue to fixate to target above chance (50%). If, on the other hand, there are novel objects as the distractors then the question might differ from one that asks 'how does A not fit A?' to one that asks 'how does A fit A-B?' (e.g., Aslin, 2007; Delle Luche et al., 2015; Houston-Price & Nakai, 2004). Overall, this procedure is an alternative method within mispronunciation studies which can achieve bias-free estimates of lexical processing by controlling for the presence of learning biases. The subsequent discussion will review the literature that has used the IPL task with novel distractors.

5.1.3 Methodological consideration with using novel distractors in IPL

Following the findings that infants can discriminate correct and incorrect pronunciations of familiar words (e.g., Mani & Plunkett, 2007), White and Morgan (2008) suggested that using familiar objects in the IPL task might have produced a mispronunciation bias. That is, when presented with a picture of a noun that is stored in an infant's lexicon such as a /cat/ alongside a familiar target image /bus/, on hearing the initial consonant MP /dus/ infants have no choice but to interpret this as a mispronunciation of the target word. In this way, the referential context could have determined the MP effect. Instead, they advised that if the distractor is not known, it will not compete as a lexical entry. They hypothesised that using novel distractors would be a better test as it will enable a demonstration of graded sensitivity to the degree of mismatch between the target word and its variant.

On this basis, White and Morgan (2008) presented 19-month-old infants with a correct pronunciation "shoe", with onset-consonant 1-feature /foo/, 2-feature /voo/ and 3-feature /goo/ MPs, together with a novel distractor image. Infants showed a linearly graded sensitivity to the degree of mismatch on a single segment. That is, as compared to the correct pronunciation, infants looked for less time at /foo/, even less for /voo/ and considerably less for /goo/. Thus, it appeared that the toddlers' behaviour was modulated by the feature overlap shared by the initial consonant MP and the correct pronunciation. This result was also extended to coda consonants by using the exact same paradigm (Ren & Morgan, 2011). Again, they tested 19-month-olds with a familiar object such as /duck/ but with a 1-feature coda-MP /dut/, 2-feature coda-MP /dud/ and 3-feature coda-MP /duz/. The pattern of results showed that infants have a graded sensitivity to varying degrees of coda mispronunciations. Overall, White et al (2008) concluded that 19-month-olds represent detail about familiar words, showing graded sensitivity to the degree of phonological mismatch between heard labels and stored representations. Thus, the authors concluded that although learners utilise their phonological sensitivities flexibly as a function of the referential context, the interpretation of a mispronunciation is dependent on the degree of mismatch.

The same paradigm was used by Mani and Plunkett (2011) to examine sensitivity to vowel MPs of familiar words in English 18- and 24-month-olds. In a study using mispronunciations of 1 to 3 features, they found that 24-month-olds, but not 18-month-olds showed a marked distinction in their sensitivity to small and large vocalic mispronunciations. Consequently, the outcome in English 18-months with vowel MPs is different to what White and Morgan (2008) showed where similar aged infants were able to discriminate between small and large consonant MPs. Whilst Mani and Plunkett (2011) queried whether this contrast suggested that consonants might be more categorically represented than vowels in early lexical development, they concluded that the graded sensitivity to vocalic contrasts can be explained by the acoustic characteristics of the mispronunciations.

Overall, the current finding combined with their earlier data (e.g., Mani et al., 2008) provides clear evidence suggesting that vowels (similarly to consonants) play a role in distinguishing lexical items which is not explained by the CV hypothesis (Nespor et al., 2003). Lastly, the contrast found between consonants (White and Morgan, 2008) and vowels (Mani and Plunkett, 2011) raises further questions about differences in the underlying representations of consonants and vowels, especially when using novel distractors. As a result, in a direct replication of 3a, we tested English 21-month-olds using the same paradigm and stimuli but with novel distractors.

5.2 Experiment 3b: Consonant and vowel MP detection in English 21-month-olds with Novel Distractors

Unexpectedly, we did not find any evidence of mispronunciation detection in Experiment 3a. Therefore, the aim of Experiment 3b is similar except this time by introducing novel distractors, we will control for the mispronunciation bias and therefore increase the measure of sensitivity (K. S. White & Morgan, 2008). Pairing the familiar targets with novel distractors (e.g., an image of a cat, paired with an image of a garlic presser) together with the correct pronunciation /cat/ or MPs /gat/, /ket/, or /kad/ will allow for a measure of graded sensitivity. Of particular interest is to explore how responses to the CPs and MPs used in Experiment 3a might be affected by the type of distractor (Mani & Plunkett, 2011; White & Morgan, 2008). Would the mispronunciation effect be larger if infants do not know the label for the other object (novel distractor), as compared to a situation where they know the distractor (a familiar object)? Indeed, studies of lexical development successfully show that toddlers are sensitive to the relationship between spoken referents and pictured objects.

Principally, in this situation infants are not able to rule out the distractor simply based on knowing it, as there is no lexical representation of an image of a garlic presser stored in their lexicon; rather, this time ruling out the mispronunciations will be dependent on the degree of mismatch between the word and the target image. Following previous IPL studies using novel distractors which show that toddlers are sensitive to variations in the size of consonant MPs (Ren & Morgan, 2011; K. S. White & Morgan, 2008) and variations in the size of vowel MPs (Mani & Plunkett, 2011), we expect that this time a sensitivity between consonant and vowel MPs might emerge. In addition, a comparison of the time course of looking time between the two experiments could reveal differences in the way the distractor knowledge might modulate target recognition (see section 5.2.3). Thus, following previous research, we expect a greater sensitivity to mispronunciations than found in Experiment 3a.

Method

This experiment 3b is an exact replication of Experiment 3a. The only difference being is that this time the distractors used were novel (see Table 5.4).

Participants

Twenty-four healthy English-learning toddlers aged 21-months (M age = 20 months and 30 days) were successfully tested (including 10 boys). The data of 21 additional infants were not included in the analyes, for either being inattentive (e.g., looking at their feet) (10), or non-completion (7), experimenter error (1) and technological failures (3).

Stimuli

The target words contained the same twenty-four monosyllabic CVC target words used in Experiment 3a. Only this time, these target words were paired with novel distractors which were taken from those used by White & Morgan

(2008). The novel items were real objects, which, except for pickle, are not included in lists of familiar words on either the infant or toddler version of the MacArthur CDI (Dale & Fenson, 1996). The full list of pairs can be found in Table 5.4. Everything else was the same as in Experiment 3a.

Table 5. 4: List of stimuli is the same as 3a except for the novel distractors

Word	Target	Distractor
Ball	bɔ:l	Doorknocker
Bath	bɑ:θ	Pickle
Bed	bɛd	Fan
Bib	bɪb	Lantern
Bin	bɪn	Padlock
Boat	bəʊt	Avocado
Book	bʊk	Paint roller
Bus	bʌs	Abacus
Cat	Kæt	Beehive
Coat	kəʊt	Bullhorn
Cot	kɒt	Trophy
Cup	cʌp	Artichoke
Dog	dɒg	Hourglass
Doll	dɒl	Accordion
Duck	dʌk	Waffle maker
Fish	fɪʃ	Shuttlecock
Foot	fʊt	Bottle opener
Hat	Hæt	Pliers
Keys	ki:z	Garlic
Leg	lɛg	Tin opener
Pen	pɛn	Horseshoe
Pig	pɪg	Pump
Sheep	ʃi:p	French horn
Sock	sɒk	Barrel

Procedure

The procedure is the same as in Experiment 3a. Only this time, the infants saw novel distractors.

Scoring

The scoring was the same as in Experiment 3a. Out of 576 trials, 86 were excluded on the basis toddlers not knowing the familiar words which left a total of 490 for the analysis. Thus, 85% of all trials were retained (82% were retained in Experiment 3a).

5.2.1 Results

A repeated measure ANOVA on PTL with Naming (pre and post) and Pronunciation (correct and incorrect) as the within-subject factors revealed no overall main effect of Pronunciation type, F (1, 23) = .50, p = .48, η_p^2 = .021, but a main effect of Naming, F (1, 23) = 34.10, p < .001, η_p^2 = .597, and a significant interaction between Naming and Pronunciation, F (1, 23) = 4.87, p = .03, η_p^2 = .175, suggesting a difference across the pronunciation conditions.

Planned comparisons were conducted to compare target looking between the pre-naming and post-naming phases between the CP and MP (all conditions collapsed), and MP in each condition. As can be seen in Table 5.5, looks to the target significantly increased from the pre-naming phase to the post-naming phase following the correct pronunciations t (23) = - 4.68, p = <.001, d= - 0.95. In contrast to Experiment 3a, looks also significantly increased from the pre-naming phase to the post-naming phase following the incorrect pronunciations t (23) = - 2.69, p =. 01, d= - 0.55.

Table 5. 5: Experiment 3b: Means (Standard Deviations) for the PTL measures for CP and MPs collapsed together.

Pronunciation	Pre	Post	Effect
CP	.53 (.15)	.68 (.13)	.15 (*0.02*)
MP	.55 (.10)	.61 (.07)	.06 (0.03)

Paired-samples t-tests demonstrated that this time, there was a significant difference in the pre- and-post PTL measures between the CPs with the onset-MPs, t (23) = -2.80, p =.01, d= - 0.57, and a significant difference between the CPs with the vowel-MPs, t (23) = 2.05, p =.05, d= - 0.41. None of the other comparisons were significantly different (all > .05). As illustrated in Figure 5.2, a graded sensitivity to the degree of mismatch can be seen.

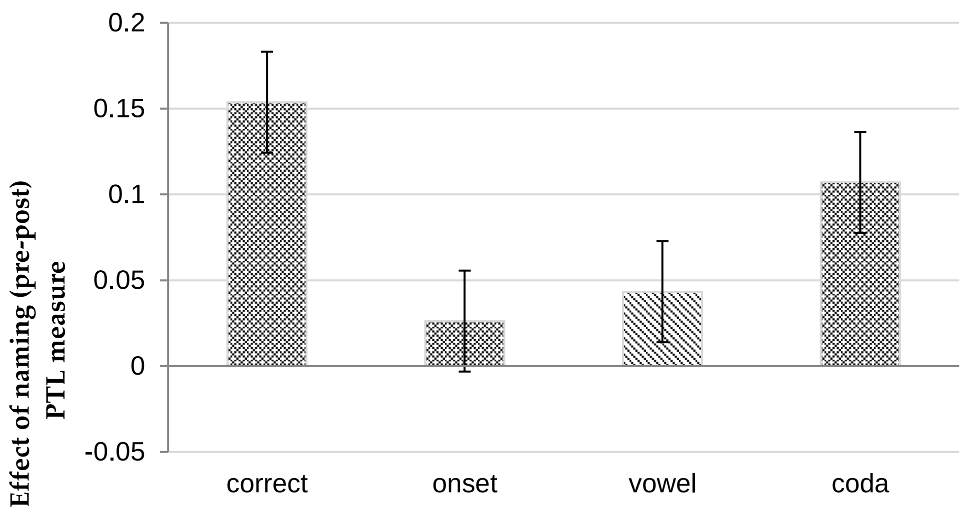

Figure 5.2: Effect of naming (pre-post) in CP, MP-Onset, MP-Vowel, and MP-Coda trials using the PTL measure. Error bars indicate the standard error.

Table 5.6 presents the PTL measures for all types of pronunciations in both the pre-naming and post-naming phases. Post hoc paired-samples *t*-tests confirmed a significant increase in target looking from the pre-post naming phase for coda-MPs, t (23) = -2.24, p =.03, d= - 0.45.

Table 5. 6: Experiment 4: Means (SD) for the PTL measures for CP and MPs separated.

Phase	CP	Onset-MP	Vowel-MP	Coda-MP
Pre-naming	.531 (.15)	.570 (.13)	.580 (.10)	.522 (.16)
Post-naming	.685 (.13)	.598 (.16)	.623 (.14)	.628 (.19)

5.2.2 Discussion of Experiment 3b

The main aim of Experiment 3b was to retest the consonant and vowel hypothesis as in Experiment 3a, by using novel distractors to increase sensitivity of the method. We hypothesised that the failure for toddlers in Experiment 3a to show an overall mispronunciation effect might be based on the presence of the competing, familiar distractor. The use of novel distractors in Experiment 3b has allowed toddlers to demonstrate that they do use their phonological sensitivities flexibly as a function of the referential context.

Across the two experiments, we found that infants looked significantly longer at the target object, but here, in contrast to Experiment 3a, they also looked significantly less at the target object when they were presented with the mispronunciations. However, this MP effect was most likely due to the consonant onset and vowel mispronunciation, as the consonant coda mispronunciation produced looking times towards the target like those of the correct pronunciation. To some extent, this pattern of results resembles what was observed in Experiment 3a (see figure 5.1), but here significant differences were observed. A visual comparison between the two experiments for post-naming PTL can be made in the time-course graphs provided in Figure 5.3 (see below).

As a final analysis, we combined the results of the two experiments and examined the time course of looking times in the different conditions. Up until now, no direct comparison between the two experimental settings (familiar distractor vs novel distractor) has been made.

5.2.3 Combined Results

Data from both Experiment 3a and Experiment 3b were pooled together to test for a main effect of Experiment or an interaction. An Independent-Samples *t*-test first confirmed that there was no significant difference between the total numbers of trials analysed between Experiment 3a (Mean: 19.7) and Experiment 3b (Mean: 20.4), t (46) = - 0.77, p = .44.

The dependent measure was the difference between PTL in the post-naming phase and the PTL in the pre-naming phase (referred to as the naming index). A repeated measure ANOVA on naming index with Pronunciation (correct and incorrect) as within-participant factors and Experiment type as the between-participant factors revealed a global effect of Pronunciation, F (3, 138) =

2.90, p = 0.03, η_p^2 = .059, no main effect of Experiment type, F (1, 46) = 1.56, p = .21, η_p^2 = .033 and no interaction between Experiment and Pronunciation, F (3, 138) = 1.28, p = .28, η_p^2 = .027.

Exploring the main effect of pronunciation, a repeated measure ANOVA with Naming (pre and post) and Pronunciation (CP, onset-MP, vowel-MP and coda-MP) as within-participant factors, and Experiment type (familiar distractors vs novel distractors) as the between-participant factor revealed a main effect of Naming, F (1, 46) = 24.36, p = <.001, η_p^2 = .346, and an interaction between Naming and Pronunciation, F (3, 138) = 3.47, p = .018, η_p^2 = .070. The interaction is due to the CP, t (47) = -4.88, p = <.001, d= -.70, and coda-MP, t (47) = -3.83, p = <.001, d= -0.55, showing a significant Naming effect whilst the onset-MP and vowel-MP do not (both = >.05). There was no interaction between Naming and Experiment, F (1, 46) = 1.77, p = .19, η_p^2 = .037. There was no triple interaction between Naming, Experiment and Distractor, F (3, 138) = .957, p = .41, η_p^2 = .020.

5.2.4 Time course plots for PTL for familiar and novel distractors

Proportion of looks as a function of time and pronunciation type were plotted for both Experiment 3a and 3b (e.g., Fernald, Zangl, Portillo, & Marchman, 2008). In order to prevent any contamination from the utterances, the analysis window began at the onset of each pronunciation type. Thus, the analysis window was aligned with the MP respectively. A visual inspection of the plots reveals that PTL is higher overall when using novel distractors compared to when using familiar distractors. As can be seen in the second graph (in Figure 5.3b) which shows the post-naming time-course for novel distractors, the onset of target recognition appears closer to that of the correct pronunciation as a function of the mismatch. That is, the MP appears to be dependent on the degree of similarity with the distractor, since the novel distractor shows that onset MP and vowel MP are detected incrementally before the coda MP compared to the correct pronunciation. This outcome is in line with White and Morgan (2008) who concluded that the interpretation of a mispronunciation in IPL tasks is dependent on the degree of mismatch with the distractor.

As seen in Figure 5.3b, the post-naming PTL for the Novel condition shows that all pronunciations are above the average of 50% of looks to the target. Infants looked significantly longer at the target object and their word-recognition is clearer. In contrast, in the first experiment (the Familiar condition) all pronunciations except the onset MPs are below the average of 50% of looks to the target. All looks to the target in the Familiar condition do not increase above 50% until 3.4 seconds into the post naming phase. Thus, word-recognition is not as clear and occurs later on in the post-naming phase. As mentioned in the discussion for Experiment 3b, to some extent, this pattern of results resembles what was observed in Experiment 3a (see figure 5.1), but here significant differences were observed. That said, whilst the distractor does contribute to clearer word recognition responses, it does not change the overall pattern of the results. Overall, we do not find a difference between onset and vowel mispronunciations, and coda changes do not block recognition, but delay it.

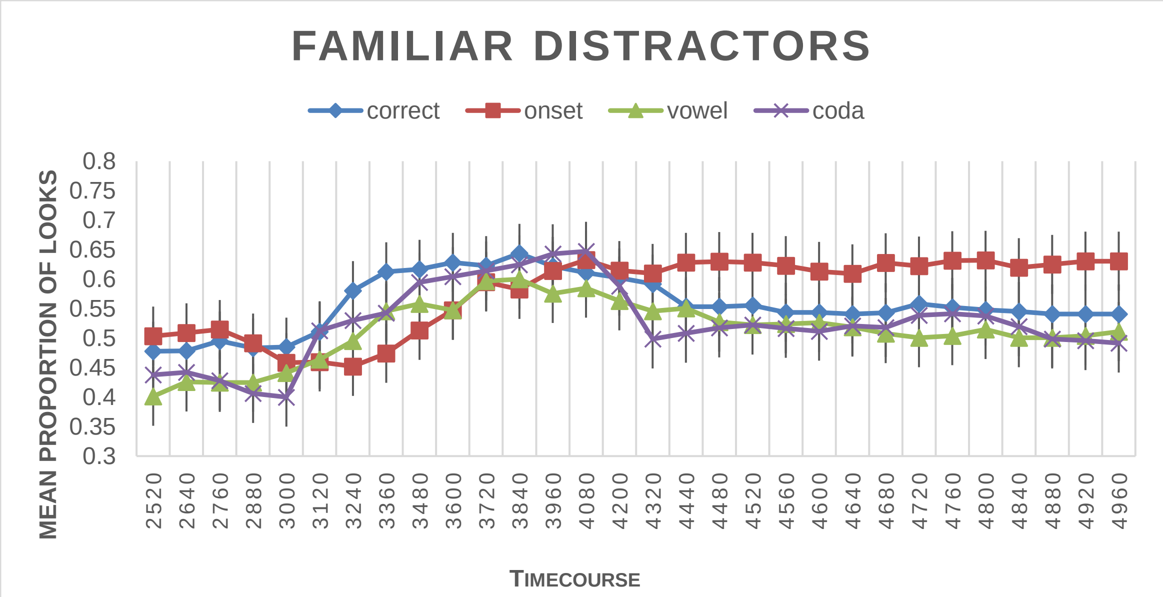

Figure 5.3a: Time Course plot (in ms, with SD) of PTL post-naming for familiar distractors.

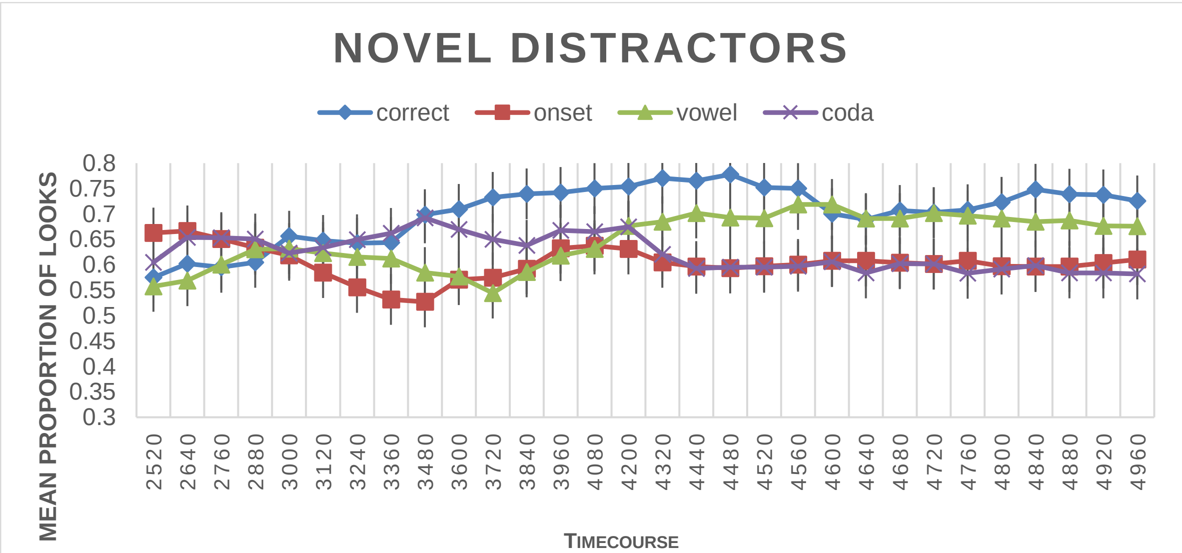

Figure 5.3b: Time Course plot (in ms, with SD) of PTL post-naming for novel distractors.

5.2.4 Discussion of Infant Experiments (3a and 3b)

The aim of this study was to examine the status of consonants and vowels in infants' lexical representations. Experiment 3a aimed to compare how consonants and vowels would affect word recognition in English infants using the standard IPL procedure. Overall, the results showed a naming effect revealing an increase in looking times in the post-naming phase. However, unexpectedly the results failed to replicate previous studies showing that infants can detect mispronunciations of familiar words (e.g., Mani & Plunkett, 2007). Thus, the next experiment (Experiment 3b) aimed to increase toddlers' sensitivity to mispronunciations by using novel objects as distractors. This led to a significant interaction between naming and pronunciation, suggesting that when using novel distractors toddlers treated correct pronunciations and mispronunciations differently, compared to when tested with familiar objects.

Taken together, whilst Experiment 3a's results do not reveal much, Experiment 3b shows a clear gradation effect between onset, vowel and coda mispronunciations. The timelines (see Figure 5.3a/b) reveal that for both experiments but clearly for Experiment 3b, infants look longer at the target post-naming in the coda-MP, but then looks are dropped considerably towards the end of the trial when the mismatch is detected. This suggests that toddlers identify the target word based on the first two segments (onset and vowel) before hearing coda-MPs (Fernald, Swingley, & Pinto, 2001; Swingley, Pinto, & Fernald, 1999). As such, our results show that within a referential context, the first portion of the word is enough for word recognition as it is often found with adults (Dahan & Tanenhaus, 2005). This shows that 21-month-olds infants are able to make use of phonetic information incrementally in a similar way as adults (e.g., McClelland & Elman, 1986), rapidly identifying spoken words before their acoustic offset as also previously found with infants at this age (Swingley, 2009a). So how can the results from this study compare to Mani and Plunkett's (2007) findings?

Firstly, using a standard IPL task, Mani and Plunkett (2007) found that unequivocally from 18-months, infants demonstrated an equal sensitivity to onset consonant and medial vowel MPs of familiar words compared to the correct pronunciations. There were however two key differences with ours and their study. For one, the proportion of trials and children tested per condition differed. In their study, 56 children completed 8 trials each, 4 of which were correct and 2 each for the onset and vowel MP, compared to ours which consisted of 24 trials per child, including 4 pronunciation types. However, the key difference is with our introduction of the coda-MP at test.

Overall, and in agreement with Mani and Plunkett (2007), we did not observe a C-bias in English toddlers. However, as a methodological advance, we have shown that when the question changes from 'how does A not fit A?' to one that asks 'how does A fit A-B?', a clear recency effect emerges indicating that an early mismatch is better detected than a later mismatch. This supports previous evidence showing that speech processing is continuous, that is, as acoustic-information is heard, children's (and adults') interpretations of speech is updated incrementally (Dahan & Tanenhaus, 2005; Swingley et al., 1999).

In conclusion, the English evidence suggests that the lexical consonant bias in word-learning emerges later in childhood than in French Spanish, or Italian (Floccia et al., 2014; Nazzi et al., 2009), favouring a language-specific explanation where the emergence of the bias depends on the distributional (e.g., Keidel, Jenison, Kluender, & Seidenberg, 2007) or acoustic-phonetic characteristics of the language (Floccia et al., 2014). By employing novel distractors in the IPL paradigm, we have also confirmed that this method is a more sensitive way to explore the degree of sensitivity of mispronunciations. To disentangle the role played by acoustic/phonological (Bouchon et al., 2015; Delle Luche et al., 2017; Floccia et al., 2014), lexical information (Keidel et al., 2007; Mayor & Plunkett, 2014), or pre-lexical information (e.g., Von Holzen et al., 2018), further research in English toddlers between 23-months and 30-months is required to establish the exact origins and linguistic nature of the English C-bias.

Chapter 6: General Discussion

The research in this study was primarily motivated by a seminal proposal which states that consonants and vowels serve functional different roles in language (Nespor et al., 2003). This CV theory predicts that whereas consonants are more informative for the lexicon, vowels serve as preferential cues for processing prosody/syntax. The goal of this research was to investigate this proposal in English adults and infants, focusing on the nature and origin of the consonantal advantage in lexical processing.

The CV hypothesis (Nespor et al., 2003) predicts a difference in how participants use consonants and vowels according to whether words or structural regularities are concealed in the speech stream. For example, in a seminal paper Toro, Nespor, Mehler, and Bonatti (2008) showed that when using an artificial language, Italian adults can use consonants to extract words and vowels to extract a structural generalisation. In addition, consonants being more important than vowels in the course of adult lexical processing has received strong support from various approaches across a number of real languages (e.g., Bonatti, Peña, Nespor, & Mehler, 2005; Creel, Aslin, & Tanenhaus, 2006; Cutler, Sebastián-Gallés, Soler-Vilageliu, & Van Ooijen, 2000; Lee, Rayner, & Pollatsek, 2001). However, the nature of the consonant bias was not so clear-cut, with uncertainty as to whether it is purely phonological, or originates from a combination of orthographic and phonological processing (e.g., see Delle Luche et al., 2014; Lupker, Perea, & Davis, 2008; New & Nazzi, 2014). One possibility that could account for the potential conflict was the level of processing being tapped into between the two different priming studies, where one task only used replaced stimuli (Delle Luche et al., 2014), and the other task used both replaced and transposed stimuli (Lupker et al., 2008). Therefore, the main purpose of the first two first experiments in this study was to examine the phonological nature of the consonant bias in adults by comparing the pattern of results from an experiment that used both transposed and replaced stimuli, but in two modalities: auditory versus written word recognition.

In terms of the developmental literature, the CV proposal has also received various support in a few languages. For example, Italian and Spanish toddlers have been found to favour vowels over consonants to learn structural regularities in the speech stream (Hochmann et al., 2017; Pons & Toro, 2010). Moreover, in regard to a consonantal advantage in lexical processing, evidence has been found in Italian (Hochmann et al., 2011) and robust empirical support has been found in French (Havy & Nazzi, 2009; Nazzi, 2005; Poltrock & Nazzi, 2015). However, whilst a C&V asymmetry had been found in early language acquisition, some languages revealed a different picture. For instance, Danish infants have demonstrated a reversed bias showing a preference for vocalic information whilst learning words (Højen & Nazzi, 2016). In addition, English toddlers have shown either an equal sensitivity to both contrasts (Floccia et al., 2014; Mani & Plunkett, 2007), or a later consonant advantage at 30-months of age (Nazzi et al., 2009). Here, a variety of possibilities could account for the discrepancy, such as the task used, phoneme position, infants' age, and their linguistic exposure. Subsequently, in experiments 3a and 3b, we have addressed some of these inconsistencies by using a mispronunciation task similar to that used in Mani and Plunkett (2007), testing for the detection of mispronunciations on vowels, onset consonant and coda consonants in English speaking infants aged 21-months.

For the adult experiments (experiments 1 and 2) presented in Chapter 3, a transposed letter and replaced letter priming paradigm was used based on the procedure used in Lupker et al.'s (2008) study. In these transposed and replaced phoneme/letter tasks, Experiment 1 was adapted to be used in the auditory domain, and Experiment 2 was a visual replication. In both experiments, adults were presented with target words like /ACADEMY/ which preceded their primes that either involved transposed consonants /ADACEMY/, replaced consonants /ABANEMY/, transposed vowels /ACEDAMY/, or replaced vowels /ACIDOMY/.

In Experiment 1, if the nature of the consonant bias is phonological, an overall advantage of processing consonants over vowels should emerge (Delle Luche et al., 2014). Indeed, the English-speaking adults tested in our study showed a significant gain of processing consonants over vowels in auditory lexical access. That is, target words preceded by primes which were obtained by modifying vowels and preserving consonants yielded faster and more accurate responses, compared to primes which modified the consonants and preserved the vowels. Thus, in line with previous adult evidence, this result confirms the observation of a facilitatory effect when the target word shared the consonants with its prime. Overall, it can be concluded that the consonantal bias in lexical processing is phonological in nature.

Experiment 2 was a visual adaptation of the exact same transposed and replaced letter paradigm used in Experiment 1, as a replication of Lupker's et al., (2008) study. If the consonant bias is purely phonological, then this time a different result should emerge. Namely, following Lupker et al. (2008), the only significant finding should be with the condition that transposed the consonants of a target word. Indeed, no overall advantage was found with consonantal over vocalic information whilst making lexical decisions in the visual domain. Rather, similar to Lupker et al. (2008) the only significant advantage was for the condition that involved transposing the consonants of the target words compared replacing them. Such a positional effect on transposed letter processing has been shown in numerous masked priming paradigms (Andrews, 1996; Lupker et al., 2008; Perea, Lupker, Kinoshita, & Lupker, 2003; Schoonbaert & Grainger, 2004; Schubert, Kinoshita, & Norris, 2017).

Taken together, the adult findings suggest that the consonant priming effect does not occur at the orthographical level but rather at the phonological and lexical levels. In support of previous online priming studies (e.g., Delle Luche et al., 2014; New & Nazzi, 2014), the adult experiments in this study provide robust confirmation of the phonological interpretation for the consonant bias in lexical processing. In favour of the CV hypothesis (Nespor et al., 2003), consonants were found to contribute more than vowels in lexical access. Overall, the cross-linguistic evidence to date suggest that the division of labour between consonants and vowels in adults is not specific to the structure of the languages. Indeed, in adulthood lexical processing being more strongly associated with consonants than with vowels has been demonstrated across 13 languages, from seven language families and in various lexically related tasks as reviewed in Chapter 2 (see also Nazzi & Cutler, (2019).

How would these results constrain models of word recognition? As discussed in the main introduction (Chapter 1), most spoken word recognition models (e.g., NAM, PARSYN, or Cohort) do not assign differences between consonants and vowels, which means similar priming effects should be observed. In regard to Experiment 1, the main finding was the facilitatory effect of the consonant related primes for lexical access in an auditory transposed and replaced lexical task. This consonant bias is not accounted for by most spoken word recognition models. For example, the PARSYN model based on neighbourhood activation (Luce et al., 2000), suggests that phonological similarity between a prime and its target word leads to inhibition. Indeed, as highlighted in Chapter 1, many studies support this hypothesis (e.g., Magnuson, Dixon, Tanenhaus, & Aslin, 2007). However, that study (and others) used real words as primes and targets. With non-word primes, the time course of activation is most likely to be different because non-words are less likely to be mistaken for real words (Delle Luche et al., 2014). Thus, activation of the target word is contained at the phonological or pre-lexical level, where activation is always facilitatory.

Of course, to make such clear theoretical statements as made in the paragraph above, one would need to simulate the experiments using these models. Although it would be very informative to simulate the experiments on these models, it is beyond the scope of the present research. Not all these models have been implemented, and when they have (e.g., TRACE: Mayor & Plunkett, 2014), it would require me to develop a whole new range of expertise and would correspond to the addition of a new experiment by itself. Besides, it is very common in experimental papers to discuss the fit between data and hypotheses drawn from models, without providing simulations, and I believe that this approach fits the level of granularity that is discussed here.

One account for the consonant bias was offered which replicated the consonant-vowel asymmetry in a TRACE model implemented on a developing lexicon (Mayor & Plunkett, 2012, 2014). As emphasised in Chapter 1, TRACE predicts that an increased sensitivity to consonant changes is related to the increasing size of cohort competitors with vocabulary size, which accommodates our behavioural data given the amount of linguistic experience participants would have accumulated by adulthood. Another way of accounting for the consonant bias which was also pointed out in the main introduction, is that phonemes might not exclusively be activated in isolation but that the skeletons or phoneme tiers (consonants) may also activate the network (Delle Luche et al., 2014).

The other main finding was the significant interaction between prime type (transposed and replaced) and phoneme type (consonants and vowels) which led us to a couple of conclusions. This interaction revealed that replacing consonants impaired lexical access more than transposing them. In contrast, replacing vowels was less damaging than transposing them. The first part of the interaction confirms that consonant transposed primes lead to more lexical activation to the target than replaced consonant primes. TRACE predicts that lexical activation that shares any similarity, including temporal locations, can occur at any point, and no mismatch inhibition is involved. Thus, this model posits that transposed consonants would lead to more lexical activation than replaced consonant primes. Subsequently, for the first part of the interaction involving consonants, again TRACE can accommodate these consonantal differences.

For the other part to the interaction, we suggested that the replaced vowel primes (e.g., ACIDOMY) might have led to more lexical activation (ACADEMY) than transposing them (ACEDAMY) due to the fact that switching vowel is often a signal of grammatical changes (e.g., Kielar, Joanisse, & Hare, 2008). That is, transposing vowels might signal to the word recognition system that the prime and target are contrasted at the grammatical level, which would activate further processing, compared to a situation involving replaced vowels, where there would be no other link than global similarity. In that way, we suggested that future research could compare consonant and vowel priming for regular and irregular verbs. Interestingly, in the visual domain the same effect has been reported whereby vowel transposed non-words such as CISANO (for CASINO), were more difficult to reject than replaced primes (CESUNO) in an un-primed lexical decision task (Perea & Lupker, 2004), though the difference was substantially smaller than between the consonant transposition and consonant replacement non-words. However, comparing non-word primes created by transposing and replacing consonants and vowels in a priming lexical decision task, often show the opposite pattern. Although the effect is not significant, replacement vowels appear marginally more difficult to reject compared to transposing vowels (Lupker et al., 2008; Perea & Lupker, 2004).

In summary, the adult evidence showing a consonant advantage favours the first part of the CV hypothesis regarding consonants being more important for lexical processing (Nespor et al., 2003). Furthermore, the present evidence provides robust support for the phonological interpretation of the consonant bias (e.g., Delle Luche et al., 2014). In terms of the vocalic feature of the CV claim, a somewhat attentive summation is that the observed priming difference between replacing and transposing vowels could reflect some aspect of morphological priming. However, to reach a more conclusive decision, future research would be required. Lastly, the fact that no consonant bias was found for non-word targets suggests that the consonant bias is related to lexical access more than pre-lexical processing.

In Experiment 2, the visual adaptation of the first experiment, we predicted that if the consonant transposed letter effect reflects more orthographic processing than phonological, then the only advantage this time should be for consonant transposed letters. This prediction was based on numerous transposed letters experiments showing that non-word primes created by transposing two nonadjacent consonants of a real word leads to enhanced lexical access as compared to replacing two nonadjacent consonants of a real word (Perea & Lupker, 2004; Schubert et al., 2017). Indeed, participants in our experiment revealed a significant consonant transposed letter effect, whereby ADACEMY led to faster and more accurate responses to its target word ACADEMY compared to a prime that replaced consonants ABANEMY. Following previous English TL studies (Lupker et al., 2008), we found no overall advantage with consonant preserved primes over vowel preserved primes.

In regard to the vowel conditions, and in contrast to what was found in the auditory TL version, our visual TL data showed similar findings to Lupker et al.'s (2008) results. That is, whilst not significant, non-word primes involving vowel transposition ACEDAMY were responded to much quicker than vowel replacements ACIDOMY. Overall, in the visual domain the impact of transposing two vowels is weaker than transposing two consonants. However, Lupker et al. (2008) favours the argument that this V-V effect can be explained in terms of the frequency of vowels and consonants in the language, and not by their functional status as such.

These visual behavioural results pose a similar challenge to models of letter position coding as spoken word recognition models, since none distinguishes between orthographic consonants and vowels (Schubert et al., 2017). A relatively new account - the CV pattern theory does consider C/V status with orthographic units, but as yet cannot account for TL similarity (Chetail et al., 2014, 2016). This theory posits that the C/V status is represented by vowel-centred units. For example, GALA would be initially represented by its constituent letters /G/ /A/ /L/ /A/, and then by two units corresponding to vowels /GA/, /LA/. Nevertheless, as discussed in the main introduction, the CV structure can be reconciled more readily with reading models such as SOLAR (Davis, 2010). Here, order is presented as an activation gradient over all of the letter in the input. Thus, Davis (2010) argues that when transposing consonant location within a word, the transitional probability is simply reversed e.g., CANISO activates its target word CASINO. In contrast, a consonant replaced prime such as CARIPO activates a whole new outcome.

Overall, the main findings of the adult experiments described in this thesis show further evidence of the phonological nature of the consonant bias in lexical processing. This main result adds to the cross-linguistic adult evidence, consistent with Nespor et al.'s (2003) hypothesis of consonants being more important than vowels for word identification, at the phonological level. Our next main aim of this thesis was to explore the developmental origins of the consonant bias. Whilst the cross-linguistic adult data mostly converge towards the observation of consonantal bias in lexical processing, the cross-linguistic developmental data is more complex.

For the infant studies presented in Chapter 5, an IPL paradigm was used based on the procedure used in Mani and Plunkett's (2007) study, in Experiment 1, and White and Morgan's (2008) in Experiment 2. In both tasks, toddlers aged 21-months were presented with familiar CVC words either correctly or incorrectly produced. The mispronunciations (MPs) occurred on the onset consonant, medial vowel or coda consonant. The mispronunciations differed by one feature where possible e.g., for the target word /bib/, the onset consonant MP was /dib/, the medial vowel MP /beb/ and coda consonant MP /bip/. In Experiment 1, in line with Mani and Plunket (2007), the target words were paired with familiar distractors (e.g., carrot), while in Experiment 2, as in White and Morgan (2008), the target words were paired with unfamiliar distractors (e.g., lantern).

The English-speaking infants tested in our study did not show any greater sensitivity to consonants over vowels in their recognition of familiar words. In Experiment 3a, whilst the results showed an overall increase in looking times in the post-naming phase (naming effect), toddlers failed to detect the mispronunciations of familiar words observed in previous studies (e.g., Mani & Plunkett, 2007). Experiment 3b used novel distractors in an attempt to increase infants' sensitivity to mispronunciations. This led to a significant interaction between naming and pronunciations, showing this time that infants treated the correct pronunciations differently to the mispronunciations. Overall, and in support of Mani and Plunkett's (2007), we did not observe a consonant bias in English toddlers.

Interestingly, in terms of positional effects between consonants and vowels in toddler word recognition, the English 21-month-olds tested in Experiment 3b were able to make use of phonetic information incrementally. In this experiment, which included novel distractors, coda changes appeared to be less well perceived than the onset changes. However, by inspecting the timelines graphs it could be seen that whilst infants looked longer at the target post-naming in the coda mispronunciation condition, looks dropped considerably towards the end of the trial as soon as the mismatch was detected. This suggests that toddlers identify familiar words based on the first two units, onset consonant and medial vowel. In contrast to French toddlers where no positional effect appears (Havy et al., 2014), a notable weakened effect was observed in English. Subsequently, our findings reveal that within a referential context, the first portion of the word is enough for toddler word recognition as has been found previously in English (Fernald et al., 2001; Swingley et al., 1999) and Dutch (Swingley, 2005) toddlers, and in the adult literature (Marslen-Wilson & Welsh, 1978; Marslen-Wilson & Zwitserlood, 1989).

6.1 Conclusion

In conclusion, the adult experiments provide further online evidence regarding the universality of the consonantal bias proposed by Nespor et al. (2003), specifically supporting the view that in spoken word recognition, consonants have a privileged role over vowels at the phonological level in English. Overall, the infant evidence suggests that the consonant bias in lexical processing emerges later in childhood in English toddlers (Floccia et al., 2014) which favours a language-specific explanation for the origins of the consonant bias, which either depends on the lexical distributional properties (Keidel et al., 2007), or the acoustic-phonetic properties of the language (Floccia et al., 2014). Future research should explore English toddlers between 23-months and 30-months to establish whether the English consonant bias might emerge earlier in familiar word recognition than reported for word learning (Floccia et al., 2014).

Appendices

Both appendices are numbered in line with the corresponding Chapter

Appendix 3A: Word targets and primes in for both Experiment 1 and 2

Word targets	**Vowel transposed**	**Consonant transposed**	**Vowel Replaced**	**Consonant replaced**
ACADEMY	acedamy	adacemy	acidomy	abanemy
ADVISORY	advosiry	adsivory	advasery	adnicory
AMATEUR	ametaur	atameur	amitour	afaneur
ANIMAL	anamil	aminal	anemol	asiral
BELOVED	belevod	bevoled	belaved	bewoted
BENEFIT	benifet	befenit	benafot	betemit
BESIDE	bisede	bedise	basude	bebine
CAFETERIA	cefateria	cateferia	cifoteria	caleberia
CAMERA	cemara	carema	cimora	casena
CAPACITY	capicaty	cacapity	capecoty	casagity
CAPITAL	capatil	catipal	capotel	cafigal
CARDINAL	cirdanal	carnidal	cordenal	carminal
CATEGORY	catogery	cagetory	catigary	capefory
CLINICAL	clinacil	clicinal	clinucel	clisimal
COGNATE	cagnote	cogtane	cugnite	cograde
COMEDY	cemody	codemy	cimudy	cobeny
CONSIDER	cinsoder	condiser	censader	conbicer
COVERAGE	covarege	corevage	covurege	cocewage
CRIMINAL	crimanil	crinimal	crimonel	crisival
DEBATE	dabete	detabe	dobute	delahe
DECADE	dacede	dedace	dicude	debave
DECIDE	dicede	dedice	dacode	detine
DELICATE	delacite	decilate	delocete	desifate
DENSITY	dinsety	dentisy	donsaty	denficy
DISPUTE	duspite	distupe	dospate	disluge
DOMINANT	domanint	donimant	domunant	docirant
EDITOR	edotir	etidor	edatur	efibor
ELABORATE	elobarate	elarobate	eluberate	elacodate
EVIDENT	evedint	edivent	evodunt	ebiwent
FORTUNE	furtone	fornute	fertane	formuke
GRATEFUL	grutefal	grafetul	grotefel	gralekul
HERITAGE	heratige	hetirage	horetage	helicage
INDICATE	indacite	incibate	inducete	insibate
LIBERAL	laberil	lirebal	laberel	linedal
LITERAL	lateril	liretal	laterel	linefal
LOCATE	lacote	lotace	lucete	lofase
LOGICAL	logacil	locigal	logecul	losipal
MARGINAL	mirganal	marnigal	morgenal	marmipal
MARINE	mirane	manire	merone	macise
MEDICINE	midecine	mecidine	madocine	mesibine
MEMORY	momery	meromy	mimary	menowy
MILITARY	milatiry	mitilary	milutery	mifikary
MISTAKE	mastike	miskate	mosteke	mishafe
MOBILE	mibole	molibe	mebale	motide

MODERATE	modarete	moredate	modurite	monebate
NUMERICAL	numirecal	nuremical	numurocal	nunewical
OPERATOR	oparetor	orepator	opirator	onegator
OPTIMAL	optamil	opmital	optamul	opcifal
ORIGINAL	origanil	orinigal	origonel	orimipal
PACIFIC	picafic	paficic	pecofic	patisic
PARENT	perant	panert	porint	pamest
POLICY	pilocy	pocily	palecy	posity
POPULAR	pupolar	polupar	pepilar	potugar
PROPOSAL	propasol	prosopal	propusil	procogal
PROVIDE	privode	prodive	pravude	probice
QUALIFY	quilafy	quafily	quelofy	quakity
QUALITY	quilaty	quatily	queloty	quafidy
RADICAL	radacil	racidal	radocel	rasibal
RAPIDLY	ripadly	radiply	repodly	rabigly
REFUSAL	refasul	resufal	refosil	renutal
REGULAR	rugelar	relugar	ragolar	retupar
RELATIVE	relitave	retalive	reletove	refakive
RELIGION	rilegion	regilion	ralugion	repifion
REMOTE	romete	retome	ramute	relone
REMOVAL	remavol	revomal	remavil	reconal
RESIDENT	resedint	redisent	resadont	rebicent
RESUME	ruseme	remuse	rasime	revune
ROMANTIC	ramontic	ronamtic	remuntic	rovastic
SALINE	silane	sanile	selone	samite
SENATOR	sanetor	setanor	sonitor	selamor
SENTIMENT	sintement	senmitent	santoment	senvilent
SPECIFIC	spicefic	speficic	spocafic	spetisic
SPECIMEN	specemin	spemicen	specuman	speniven
STOLEN	stelon	sloten	stalan	skofen
STRATEGY	stretagy	stragety	strotigy	strapely
TRAGEDY	tregady	tradegy	trigody	trakety
TRIBUTE	trubite	tritube	trabete	trilude
VALIDITY	viladity	vadility	voledity	vabifity
VELOCITY	velicoty	vecolity	velecaty	vesofity
VETERAN	veteran	veretan	vatiran	vecelan

Appendix 3B: Lexical characteristics of the target words and their primes

Variable	Target word	CT	VT	CR	VR
Number of syllables	2.9 (0.6)	same	same	same	same
Phonological uniqueness	3.6 (2.7)	4.2 (0.8)	4.2 (1.0)	4.2 (1.0)	4.0 (0.7)
Orthographical uniqueness	1.9 (2.9)	4.6 (0.8)	4.5 (0.8)	4.5 (0.8)	4.5 (0.8)
Phon. Neighbourhood	4.5 (12.1)	0.4 (2.0)	1.2 (5.3)	0.1 (0.3)	0.5 (2.0)
Orth. neighbourhood	3.9 (26.5)	0.6 (2.9)	0.2 (1.4)	0.1 (0.3)	0.5 (2.4)
Phonological levenshtein (PLD20)	2.7 (0.5)	3.0 (0.5)	3.1 (0.6)	3.0 (0.6)	3.3 (0.6)
Orthographical levenshtein (OLD20)	2.6 (0.4)	2.8 (0.4)	3.0 (0.4)	2.9 (0.4)	3.0 (0.2)
Consonant Neighbour	7.9 (13.5)	4.3 (7.3)	7.9 (13.4)	6.1 (12.1)	7.9 (13.5)
Vowel Neighbour	88.9 (88.6)	88.5 (80.8)	114.2 (202.5)	88.1 (81.3)	63.5 (92.8)

References

Acha, J., & Perea, M. (2010). On the role of consonants and vowels in visual-word processing: Evidence with a letter search paradigm. *Language and Cognitive Processes*, *25*(3), 423–438. https://doi.org/10.1080/01690960903411666

Andics, A. (2006). Distinguishing between prelexical levels in speech perception: An adaptation-fMRI study. *From the Programme Director*, 47.

Andrews, S. (1996). Lexical retrieval and selection processes: Effects of transposed-letter confusability. *Journal of Memory and Language*, *35*(6), 775–800.

Arvanti, A. (2009). Rhythm, Timing and the Timing of Rhythm. *Phonetica, 66,* (1-2), 46-63.

Aslin, R., & Fiser, J. (2005). Methodological challenges for understanding cognitive development in infants. *Trends in Cognitive Sciences*, *9*(3), 92–98. https://doi.org/10.1016/j.tics.2005.01.003

Aslin, R. N. (2007). What's in a look? *Developmental Science*, *10*(1), 48–53. https://doi.org/10.1111/j.1467-7687.2007.00563.x

Baayen, R. H., Piepenbrock, R., & Gulikers, R. (1996). *Baayen: The CELEX lexical database (cd-rom)—Google Scholar*. https://scholar.google.com/scholar_lookup?title=The%20CELEX%20lexical%20database&author=R.H.%20Baayen&publication_year=1995

Berent, I., & Perfetti, C. A. (1995). A rose is a REEZ: The two-cycles model of phonology assembly in reading English. *Psychological Review*, *102*(1), 146–184. https://doi.org/10.1037/0033-295X.102.1.146

Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P. W., Kennedy, L. J., & Mehler, J. (1988). An investigation of young infants' perceptual representations of speech sounds. *Journal of Experimental Psychology: General*, *117*(1), 21.

Bertoncini, J., Floccia, C., Nazzi, T., & Mehler, J. (1995). Morae and syllables: Rhythmical basis of speech representations in neonates. *Language and Speech*, *38 (Pt 4)*, 311–329. https://doi.org/10.1177/002383099503800401

Bertoncini, Josiane, & Mehler, J. (1981). Syllables as units in infant speech perception. *Infant Behavior & Development - INFANT BEHAV DEVELOP*, *4*, 247–260. https://doi.org/10.1016/S0163-6383(81)80027-6

Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations. *Psychological Science*, *16*(6), 451.

Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2007). On consonants, vowels, chickens, and eggs. *Psychological Science*, *18*(10), 924–925.

Bouchon, C., Floccia, C., Fux, T., Adda-Decker, M., & Nazzi, T. (2015). Call me Alix, not Elix: Vowels are more important than consonants in own-name recognition at 5 months. *Developmental Science*, *18*(4), 587–598. https://doi.org/10.1111/desc.12242

Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language*, *55*(4), 627–652. https://doi.org/10.1016/j.jml.2006.08.005

Butler, J., Floccia, C., Goslin, J., & Panneton, R. (2011). Infants' Discrimination of Familiar and Unfamiliar Accents in Speech: INFANT ACCENT PERCEPTION. *Infancy*, *16*(4), 392–417. https://doi.org/10.1111/j.1532-7078.2010.00050.x

Caramazza, A., Chialant, D., Capasso, R., & Miceli, G. (2000). Separable processing of consonants and vowels. *Nature*, *403*(6768), 428. https://doi.org/10.1038/35000206

Carreiras, M., Duñabeitia, J. A., & Molinaro, N. (2009). Consonants and vowels contribute differently to visual word recognition: ERPs of relative position priming. *Cerebral Cortex*, *19*(11), 2659–2670.

Carreiras, M., Ferrand, L., Grainger, J., & Perea, M. (2005). Sequential Effects of Phonological Priming in Visual Word Recognition. *Psychological Science*, *16*(8), 585–589. https://doi.org/10.1111/j.1467-9280.2005.01579.x

Carreiras, M., Gillon-Dowens, M., Vergara, M., & Perea, M. (2008). Are vowels and consonants processed differently? Event-related potential evidence with a delayed letter paradigm. *Journal of Cognitive Neuroscience*, *21*(2), 275–288.

Carreiras, Manuel, Vergara, M., & Perea, M. (2009). ERP correlates of transposed-letter priming effects: The role of vowels versus consonants. *Psychophysiology*, *46*(1), 34–42. https://doi.org/10.1111/j.1469-8986.2008.00725.x

Chetail, F., Drabs, V., & Content, A. (2014). The role of consonant/vowel organization in perceptual discrimination. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 938–961. https://doi.org/10.1037/a0036166

Chetail, F., Treiman, R., & Content, A. (2016). Effect of consonant/vowel letter organisation on the syllable counting task: Evidence from English. *Journal of Cognitive Psychology*, *28*(1), 32–43. https://doi.org/10.1080/20445911.2015.1074582

Coath, M., Brader, J. M., Fusi, S., & Denham, S. L. (2005). Multiple views of the response of an ensemble of spectro-temporal features support concurrent classification of utterance, prosody, sex and speaker identity. *Network: Computation in Neural Systems*, *16*(2–3), 285–300. https://doi.org/10.1080/09548980500290120

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256.

Cornell, S. A., Lahiri, A., & Eulitz, C. (2013). Inequality across consonantal contrasts in speech perception: Evidence from mismatch negativity. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 757–772. https://doi.org/10.1037/a0030862

Creel, S. C., Aslin, R. N., & Tanenhaus, M. K. (2006). Acquiring an artificial lexicon: Segment type and order information in early lexical entries. *Journal of Memory and Language*, *54*(1), 1–19. https://doi.org/10.1016/j.jml.2005.09.003

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.

Cutler, Anne, & Otake, T. (2002). Rhythmic Categories in Spoken-Word Recognition. *Journal of Memory and Language*, *46*(2), 296–322. https://doi.org/10.1006/jmla.2001.2814

Cutler, Anne, Van Ooijen, B., & Norris, D. (1999). Vowels, consonants, lexical activation. *Proceedings of the Fourteenth International Congress of Phonetic Sciences*, *3*, 2053–2056.

Dahan, D., & Tanenhaus, M. K. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychonomic Bulletin & Review*, *12*(3), 453–459. https://doi.org/10.3758/BF03193787

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, *28*(1), 125–127. https://doi.org/10.3758/BF03203646

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, *117*(3), 713–758. https://doi.org/10.1037/a0019738

Davis, C. J., & Lupker, S. J. (2006). Masked inhibitory priming in English: Evidence for lexical inhibition. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(3), 668–687. https://doi.org/10.1037/0096-1523.32.3.668

Davis, M. H. (2000). Lexical segmentation in spoken word recognition. *Unpublished PhD Thesis, Birkbeck College, University of London. Available: Http://Www. Mrccbu. Cam. Ac. Uk/Personal/Matt. Davis/Thesis/Index. Html*.

Dehaene-Lambertz, G., & Baillet, S. (1998). A phonological representation in the infant brain: *NeuroReport*, *9*(8), 1885–1888. https://doi.org/10.1097/00001756-199806010-00040

Delle Luche, C., Durrant, S., Poltrock, S., & Floccia, C. (2015). A methodological investigation of the Intermodal Preferential Looking paradigm: Methods of analyses, picture selection and data rejection criteria. *Infant Behavior and Development*, *40*, 151–172. https://doi.org/10.1016/j.infbeh.2015.05.005

Delle Luche, C., Floccia, C., Granjon, L., & Nazzi, T. (2017). Infants' First Words are not Phonetically Specified: Own Name Recognition in British English-Learning 5-Month-Olds. *Infancy*, *22*(3), 362–388. https://doi.org/10.1111/infa.12151

Delle Luche, C., Poltrock, S., Goslin, J., New, B., Floccia, C., & Nazzi, T. (2014). Differential processing of consonants and vowels in the auditory modality: A cross-linguistic study. *Journal of Memory and Language*, *72*, 1–15. https://doi.org/10.1016/j.jml.2013.12.001

Dimitropoulou, M., Duñabeitia, J. A., & Carreiras, M. (2011). Phonology by itself: Masked phonological priming effects with and without orthographic overlap. *Journal of Cognitive Psychology*, *23*(2), 185–203. https://doi.org/10.1080/20445911.2011.477811

Duñabeitia, J. A., & Carreiras, M. (2011). The relative position priming effect depends on whether letters are vowels or consonants. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1143.

Dunabeitia, J. A., & Molinaro, N. (Eds.). (2014). *At the doors of lexical access: The importance of the first 250 milliseconds in reading*. Frontiers Media SA. http://www.frontiersin.org/books/At_the_doors_of_lexical_access_The_importance_of_the_first_250_milliseconds_in_reading/318

Eimas, P. D. (1999). Segmental and syllabic representations in the perception of speech by young infants. *The Journal of the Acoustical Society of America*, *105*(3), 1901–1911. https://doi.org/10.1121/1.426726

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*(2), 179–211. https://doi.org/10.1016/0364-0213(90)90002-E

Endress, A. D., & Bonatti, L. L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition*, *105*(2), 247–299.

Endress, Ansgar D., & Bonatti, L. L. (2016). Words, rules, and mechanisms of language acquisition: Mechanisms of language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *7*(1), 19–35. https://doi.org/10.1002/wcs.1376

Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language: Theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review*, *37*, 66–108. https://doi.org/10.1016/j.dr.2015.05.002

Escudero, P., Mulak, K. E., & Vlach, H. (2016). Cross-situational learning of minimal word pairs. *Cognitive Science*, *40*(2), 455–465. https://doi.org/10.1111/cogs.12243

Fernald, A. E., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Language Acquisition and Language Disorders* (Vol. 44, pp. 97–135). John Benjamins Publishing Company. https://benjamins.com/catalog/lald.44.06fer

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, *72*(4), 1003–1015.

Ferrand, L., & Grainger, J. (1993). The time course of orthographic and phonological code activation in the early phases of visual word recognition. *Bulletin of the Psychonomic Society*, *31*(2), 119–122.

Floccia, C., Kolinsky, R., Dodane, C., & Morais, J. (2003). Discriminating spoken words in French: The role of the syllable and the CV phonological skeleton. *Language* *and Cognitive Processes*, *18*(3), 241–267. https://doi.org/10.1080/01690960244000009

Floccia, C., Nazzi, T., Delle Luche, C., Poltrock, S., & Goslin, J. (2014). English-learning one- to two-year-olds do not show a consonant bias in word learning. *Journal of Child Language*, *41*(05), 1085–1114. https://doi.org/10.1017/S0305000913000287

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 sentencesa). *The Journal of the Acoustical Society of America*, *131*(2), 1490–1501.

Foss, D. J., & Swinney, D. A. (1973). On the psychological reality of the phoneme: Perception, identification, and consciousness. *Journal of Verbal Learning & Verbal Behavior*, *12*(3), 246–257. https://doi.org/10.1016/S0022-5371(73)80069-6

Frauenfelder, U. H., & Floccia, C. (1999). The Recognition of Spoken Words. In *Language Comprehension: A Biological Perspective* (pp. 1–40). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-59967-5_1

Frost, R. (1998). Toward a Strong Phonological Theory of Visual Word Recognition: True Issues and False Trails. *Psychological Bulletin*, *123*(1), 71–79.

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming Lexical Neighbors of Spoken Words: Effects of Competition and Inhibition. *Journal of Memory and Language*, *28*(5), 501–518.

Goldsmith, J. A. (1995). Phonological theory. *The Handbook of Phonological Theory*, *1*, 23.

Golinkoff, R. M., Ma, W., Song, L., & Hirsh-Pasek, K. (2013). Twenty-Five Years Using the Intermodal Preferential Looking Paradigm to Study Language Acquisition: What Have We Learned? *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *8*(3), 316–339. https://doi.org/10.1177/1745691613484936

Gómez, D. M., Mok, P., Ordin, M., Mehler, J., & Nespor, M. (2018). Statistical Speech Segmentation in Tone Languages: The Role of Lexical Tones. *Language and Speech*, *61*(1), 84–96. https://doi.org/10.1177/0023830917706529

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review*, *115*(3), 577–600. https://doi.org/10.1037/a0012667

Gonzalez-Gomez, N., Poltrock, S., & Nazzi, T. (2013). A “Bat” Is Easier to Learn than a “Tab”: Effects of Relative Phonotactic Frequency on Infant Word Learning. *PLoS ONE*, *8*(3), e59601. https://doi.org/10.1371/journal.pone.0059601

Goslin, J., & Frauenfelder, U. H. (2001). A Comparison of Theoretical and Human Syllabification. *Language and Speech*, *44*(4), 409–436. https://doi.org/10.1177/00238309010440040101

Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes*, *23*(1), 1–35. https://doi.org/10.1080/01690960701578013

Grainger, J., & Ferrand, L. (1996). Masked orthographic and phonological priming in visual word recognition and naming: Cross-task comparisons. *Journal of Memory and Language*, *35*(5), 623–647.

Grainger, J., Kiyonaga, K., & Holcomb, P. J. (2006). The Time Course of Orthographic and Phonological Code Activation. *Psychological Science*, *17*(12), 1021–1026. https://doi.org/10.1111/j.1467-9280.2006.01821.x

Grainger, J., & Whitney, C. (2004). Does the huamn mnid raed wrods as a wlohe? *Trends in Cognitive Sciences*, *8*(2), 58–59. https://doi.org/10.1016/j.tics.2003.11.006

Gregová, R. (2010). A Comparative analysis of consonant clusters in English and in Slovak. *Bulletin of the Transilavania Univeristy of Brasov*, *3*(52), 7.

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language*, *27*(3), 689–705. https://doi.org/10.1017/S0305000900004414

Havy, & Nazzi, T. (2009). Better Processing of Consonantal Over Vocalic Information in Word Learning at 16 Months of Age. *Infancy*, *14*(4), 439–456. https://doi.org/10.1080/15250000902996532

Havy, Mélanie, 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.

Hochmann, J.-R., Benavides-Varela, S., Fló, A., Nespor, M., & Mehler, J. (2017). Bias for Vocalic Over Consonantal Information in 6-Month-Olds. *Infancy*. https://doi.org/10.1111/infa.12203

Hochmann, J.-R., Benavides-Varela, S., Nespor, M., & Mehler, J. (2011). Consonants and vowels: Different roles in early language acquisition: Consonants and vowels in language acquisition. *Developmental Science*, *14*(6), 1445–1458. https://doi.org/10.1111/j.1467-7687.2011.01089.x

Højen, A., & Nazzi, T. (2016). Vowel bias in Danish word-learning: Processing biases are language-specific. *Developmental Science*, *19*(1), 41–49. https://doi.org/10.1111/desc.12286

Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, *13*(4), 341–348. https://doi.org/10.1002/icd.364

Jared, D., & Seidenberg, M. S. (1991). Does word identification proceed from spelling to sound to meaning. *Journal of Experimental Psychology: General*, *120*, 358–394.

Keidel, J. L., Jenison, R. L., Kluender, K. R., & Seidenberg, M. S. (2007). Does grammar constrain statistical learning? Commentary on Bonatti, Peña, Nespor, and Mehler (2005). *Psychological Science*, *18*(10), 922–923.

Kielar, A., Joanisse, M. F., & Hare, M. L. (2008). Priming English past tense verbs: Rules or statistics? *Journal of Memory and Language*, *58*(2), 327–346. https://doi.org/10.1016/j.jml.2007.10.002

Kinoshita, S., & Norris, D. (2013). Letter order is not coded by open bigrams. *Journal of Memory and Language*, *69*(2), 135–150. https://doi.org/10.1016/j.jml.2013.03.003

Klatt, D. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, *7*, 279–312.

Laakso, A., & Calvo, P. (2011). How Many Mechanisms Are Needed to Analyze Speech? A Connectionist Simulation of Structural Rule Learning in Artificial Language Acquisition. *Cognitive Science*, *35*(7), 1243–1281. https://doi.org/10.1111/j.1551-6709.2011.01191.x

Laakso, A., & Calvo, P. (2008). A connectionist simulation of structural rule learning in language acquisition. *Proceedings of the Cognitive Science Society*, *30*.

Ladefoged, P., & Disner, S. F. (2012). *Vowels and Consonants* (3rd Revised edition edition). Wiley-Blackwell.

Lee, H.-W., Rayner, K., & Pollatsek, A. (2001). The Relative Contribution of Consonants and Vowels to Word Identification during Reading. *Journal of Memory and Language*, *44*(2), 189–205. https://doi.org/10.1006/jmla.2000.2725

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the Speech Code. *Psychological Review*, *74*(6), 431–461.

Luce, P. A., Goldinger, S. D., Auer, E. T., & Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception & Psychophysics*, *62*(3), 615–625. https://doi.org/10.3758/BF03212113

Luce, P. A., & Pisoni, D. B. (1998). Recognizing Spoken Words: The Neighborhood Activation Model. *Ear and Hearing*, *19*(1), 1–36.

Lupker, S. J., Perea, M., & Davis, C. J. (2008). Transposed-letter effects: Consonants, vowels and letter frequency. *Language and Cognitive Processes*, *23*(1), 93–116.

Maddieson, I., Bhattacharya, T., Smith, D. E., & Croft, W. (2011). Geographical distribution of phonological complexity. *Linguistic Typology*, *15*(2). https://doi.org/10.1515/lity.2011.020

Magnuson, J. S., Dixon, J. A., Tanenhaus, M. K., & Aslin, R. N. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science*, *31*(1), 133–156. https://doi.org/10.1080/03640210709336987

Magnuson, J. S., Mirman, D., Luthra, S., Strauss, T., & Harris, H. D. (2018). Interaction in Spoken Word Recognition Models: Feedback Helps. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00369

Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (2003). The time course of spoken word learning and recognition: Studies with artificial lexicons. *Journal of Experimental Psychology: General*, *132*(2), 202–227. https://doi.org/10.1037/0096-3445.132.2.202

Mani, N., Coleman, J., & Plunkett, K. (2008). Phonological Specificity of Vowel Contrasts at 18-months. *Language and Speech*, *51*(1–2), 3–21. https://doi.org/10.1177/00238309080510010201

Mani, N, & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, *57*(2), 252–272. https://doi.org/10.1016/j.jml.2007.03.005

Mani, Nivedita, & Plunkett, K. (2010). Twelve-Month-Olds Know Their Cups From Their Keps and Tups: TWELVE-MONTH-OLDS KNOW THEIR CUPS FROM THEIR KEPS AND TUPS. *Infancy*, *15*(5), 445–470. https://doi.org/10.1111/j.1532-7078.2009.00027.x

Mani, Nivedita, & Plunkett, K. (2011). Does size matter? Subsegmental cues to vowel mispronunciation detection. *Journal of Child Language*, *38*(3), 606–627. https://doi.org/10.1017/S0305000910000243

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule Learning by Seven-Month-Old Infants. *Science*, *283*(5398), 77–80. https://doi.org/10.1126/science.283.5398.77

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*(1), 29–63. https://doi.org/10.1016/0010-0285(78)90018-X

Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review*, *101*(4), 653–675. https://doi.org/10.1037/0033-295X.101.4.653

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.

Mattys, S. L., & Melhorn, J. F. (2005). How do Syllables Contribute to the Perception of Spoken English? Insight from the Migration Paradigm ,
How do Syllables Contribute to the Perception of Spoken English? Insight from the Migration Paradigm. *Language and Speech*, *48*(2), 223–252. https://doi.org/10.1177/00238309050480020501

Mayor, J., & Plunkett, K. (2012). *Going with TRACE beyond Infant Mispronunciation Studies: Lexical Networks and Phoneme Competition*.

Mayor, J., & Plunkett, K. (2014). Infant word recognition: Insights from TRACE simulations. *Journal of Memory and Language*, *71*(1), 89–123. https://doi.org/10.1016/j.jml.2013.09.009

McCarthy, J. J. (1985). *Formal problems in Semitic phonology and morphology*.

McCarthy, J. J. (1988). Feature Geometry and Dependency: A Review. *Phonetica*, *45*(2–4), 84–108. https://doi.org/10.1159/000261820

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86. https://doi.org/10.1016/0010-0285(86)90015-0

McClelland, J. L., & Patterson, K. (2002). Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences*, *6*(11), 465–472.

McQueen, J. M., Dahan, D., & Cutler, A. (2003). Continuity and gradedness in speech processing. In *Phonetics and Phonology in Language Comprehension and*

Production: Differences and Similarities (pp. 39–78). Mouton de Gruyter. http://researchdirect.westernsydney.edu.au/islandora/object/uws%3A3002/

Mehler, J., Dommergues, J. Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, *20*(3), 298–305.

Mehler, J., Peña, M., Nespor, M., & Bonatti, L. (2006). SPECIAL ISSUE: ORIGINAL ARTICLE THE “SOUL” OF LANGUAGE DOES NOT USE STATISTICS: REFLECTIONS ON VOWELS AND CONSONANTS. *Cortex*, *42*, 846–854.

Miller, J. (1991). Reaction time analysis with outlier exclusion: Bias varies with sample size. The Quarterly Journal of Experimental Psychology, 43(4), 907-912. https://doi.org/10.1080%2F14640749108400962

Moates, D. R., & Marks, E. A. (2012). Vowel Mutability in Print in English and Spanish. *The Mental Lexicon*, *7*(3), 326–350. https://doi.org/10.1075/ml.7.3.04moa

Mueller, J. L., Bahlmann, J., & Friederici, A. D. (2008). The Role of Pause Cues in Language Learning: The Emergence of Event-related Potentials Related to Sequence Processing. *Journal of Cognitive Neuroscience*, *20*(5), 892–905. https://doi.org/10.1162/jocn.2008.20511

Nakajima, Y., Ueda, K., Fujimaru, S., & Ohsaka, Y. (2013). *Sonority in British English*. 025019–025019. https://doi.org/10.1121/1.4800164

Nazzi, T., & Bertoncini, J. (2009). Phonetic Specificity in Early Lexical Acquisition: New Evidence from Consonants in Coda Positions. *Language and Speech*, *52*(4), 463–480. https://doi.org/10.1177/0023830909336584

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition*, *98*(1), 13–30. https://doi.org/10.1016/j.cognition.2004.10.005

Nazzi, T. Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 756.

Nazzi, T., & Cutler, A. (2019). HOW CONSONANTS AND VOWELS SHAPE SPOKEN-LANGUAGE RECOGNITION. *Annual Review of Linguistics*, 5 (1). Available at SSRN: https://ssrn.com/abstract=3335695 or http://dx.doi.org/10.1146/annurev-linguistics-011718-011919

Nazzi, T., & Cutler, A. (2019). *How Consonants and Vowels Shape Spoken-Language Recognition* (SSRN Scholarly Paper ID 3335695). Social Science Research Network. https://papers.ssrn.com/abstract=3335695

Nazzi, T., & Floccia, C., Moquet, B., & Butler, J. (2009). Bias for consonantal information over vocalic information in 30-month-olds: Cross-linguistic evidence from French and English. *Journal of Experimental Child Psychology*, *102*(4), 522–537. https://doi.org/10.1016/j.jecp.2008.05.003

Nazzi, T., & Gopnik, A. (2001). Linguistic and cognitive abilities in infancy: When does language become a tool for categorization? *Cognition*, *80*(3), B11–B20.

Nazzi, T., & New, B. (2007). Beyond stop consonants: Consonantal specificity in early lexical acquisition. *Cognitive Development*, *22*(2), 271–279. https://doi.org/10.1016/j.cogdev.2006.10.007

Nazzi, T., Poltrock, S., & Von Holzen, K. (2016). The Developmental Origins of the Consonant Bias in Lexical Processing. *Current Directions in Psychological Science*, *25*(4), 291–296.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*(3), 226–254. https://doi.org/10.1037/0096-3445.106.3.226

Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio*, *2*(2), 203–230.

New, B., Araújo, V., & Nazzi, T. (2008). Differential processing of consonants and vowels in lexical access through reading. *Psychological Science*, *19*(12), 1223–1227.

New, B., & Nazzi, T. (2014). The time course of consonant and vowel processing during word recognition. *Language, Cognition and Neuroscience*, *29*(2), 147–157. https://doi.org/10.1080/01690965.2012.735678

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*(2), 127–162. https://doi.org/10.1016/S0010-0285(03)00128-2

Nishibayashi, L.-L., & Nazzi, T. (2016). Vowels, then consonants: Early bias switch in recognizing segmented word forms. *Cognition*, *155*, 188–203. https://doi.org/10.1016/j.cognition.2016.07.003

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*(3), 189–234. https://doi.org/10.1016/0010-0277(94)90043-4

Norris, D., McQueen, J. M., & Cutler, A. (2000). Feedback on feedback on feedback: It's feedforward. *Behavioral and Brain Sciences*, *23*(3), 352–363. https://doi.org/10.1017/S0140525X0053324X

Onnis, L., Monaghan, P., Richmond, K., & Chater, N. (2005). *Phonology impacts segmentation in online speech processing*. https://doi.org/10.1016/j.jml.2005.02.011

Parmentier, F. B. R., Turner, J., & Perez, L. (2014). A dual contribution to the involuntary semantic processing of unexpected spoken words. *Journal of Experimental Psychology: General*, *143*(1), 38–45. https://doi.org/10.1037/a0031550

Pastizzo, M. J., & Feldman, L. B. (2002). Discrepancies between orthographic and unrelated baselines in masked priming undermine a decompositional account of morphological facilitation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 244–249. https://doi.org/10.1037/0278-7393.28.1.244

Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science (New York, N.Y.)*, *298*(5593), 604–607. https://doi.org/10.1126/science.1072901

Perea, M., & Lupker, S. J. (2003). Does jugde activate COURT? Transposed-letter similarity effects in masked associative priming. *Memory & Cognition*, *31*(6), 829–841.

Perea, M., Lupker, S. J., Kinoshita, S., & Lupker, S. J. (2003). Transposed-letter confusability effects in masked form priming. *Masked Priming: State of the Art*, 97–120.

Perea, Manuel, & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language*, *51*(2), 231–246. https://doi.org/10.1016/j.jml.2004.05.005

Perea, Manuel, Marcet, A., & Acha, J. (2018). Does consonant–vowel skeletal structure play a role early in lexical processing? Evidence from masked priming. *Applied Psycholinguistics*, *39*(01), 169–186. https://doi.org/10.1017/S0142716417000431

Perea, Manuel, Moret-Tatay, C., & Carreiras, M. (2011). Facilitation versus Inhibition in the Masked Priming Same–Different Matching Task. *Quarterly Journal of Experimental Psychology*, *64*(10), 2065–2079. https://doi.org/10.1080/17470218.2011.582131

Pisoni, D. B., & Luce, P. A. (1987). Acoustic-phonetic representations in word recognition. *Cognition*, *25*(1–2), 21–52.

Pollatsek, A., Perea, M., & Carreiras, M. (2005). Doesconal prime canal more thancinal? Masked phonological priming effects in Spanish with the lexical decision task. *Memory & Cognition*, *33*(3), 557–565. https://doi.org/10.3758/BF03193071

Poltrock, S., & Nazzi, T. (2015). Consonant/vowel asymmetry in early word form recognition. *Journal of Experimental Child Psychology*, *131*, 135–148. https://doi.org/10.1016/j.jecp.2014.11.011

Pons, F., & Toro, J. M. (2010). Structural generalizations over consonants and vowels in 11-month-old infants. *Cognition*, *116*(3), 361–367. https://doi.org/10.1016/j.cognition.2010.05.013

Radeau, M., Morais, J., & Dewier, A. (1989). Phonological priming in spoken word recognition: Task effects. *Memory & Cognition*, *17*(5), 525–535.

Radeau, M., Morais, J., & Segui, J. (1995). Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(6), 1297–1311. https://doi.org/10.1037/0096-1523.21.6.1297

Ramon-Casas, M., Swingley, D., Sebastián-Gallés, N., & Bosch, L. (2009). Vowel categorization during word recognition in bilingual toddlers. *Cognitive Psychology*, *59*(1), 96–121. https://doi.org/10.1016/j.cogpsych.2009.02.002

Rastle, K., & Brysbaert, M. (2006). Masked phonological priming effects in English: Are they real? Do they matter? *Cognitive Psychology*, *53*(2), 97–145.

Rastle, K., Davis, M. H., Marslen-Wilson, W. D., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes*, *15*(4–5), 507–537. https://doi.org/10.1080/01690960050119689

Ren, J., & Morgan, J. L. (2011). Sub-segmental details in early lexical representation of consonants. *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII)*. http://titan.cog.brown.edu/research/infant_lab/Ren_sub%20segmental%20details%20early%20lexical.pdf

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(6), 906–914. https://doi.org/10.1002/wcs.78

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior*, *10*(6), 645–657. https://doi.org/10.1016/S0022-5371(71)80071-3

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. MIT Press.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Schoonbaert, S., & Grainger, J. (2004). Letter position coding in printed word perception: Effects of repeated and transposed letters. *Language and Cognitive Processes*, *19*(3), 333–367. https://doi.org/10.1080/01690960344000198

Schubert, T., Kinoshita, S., & Norris, D. (2017). What Causes the Greater Perceived Similarity of Consonant-Transposed Nonwords? *Quarterly Journal of Experimental Psychology*, 17470218.2016.1. https://doi.org/10.1080/17470218.2016.1271444

Segui, J., Frauenfelder, U., & Mehler, J. (1981). Phoneme monitoring, syllable monitoring and lexical access. *British Journal of Psychology*, *72*(4), 471–477. https://doi.org/10.1111/j.2044-8295.1981.tb01776.x

Seidenberg, M. S. (2005). Connectionist Models of Word Reading. *Current Directions in Psychological Science*, *14*(5), 238–242. https://doi.org/10.1111/j.0963-7214.2005.00372.x

Seidenberg, M. S., MacDonald, M. C., & Saffran, J. R. (2002). Does Grammar Start Where Statistics Stop? *Science*, *298*(5593), 553–554. https://doi.org/10.1126/science.1078094

Seidenberg, M. S., & McClelland, J. L. (1989). A Distributed, Developmental Model of Word Recognition and Naming. *Psychological Review*, *96*(4), 523–568.

Sharp, D., Scott, S., Cutler, A., & Wise, R. (2005). Lexical retrieval constrained by sound structure: The role of the left inferior frontal gyrus. *Brain and Language*, *92*(3), 309–319. https://doi.org/10.1016/j.bandl.2004.07.002

Soares, A. P., Perea, M., & Comesaña, M. (2014). Tracking the Emergence of the Consonant Bias in Visual-Word Recognition: Evidence with Developing Readers. *PLoS ONE*, *9*(2), e88580. https://doi.org/10.1371/journal.pone.0088580

Sussman, H. M., McCaffrey, H. A., & Matthews, S. A. (1991). An investigation of locus equations as a source of relational invariance for stop place categorization. *The Journal of the Acoustical Society of America*, *90*(3), 1309–1325. https://doi.org/10.1121/1.401923

Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, *8*(5), 432–443.

Swingley, D. (2009a). Onsets and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language*, *60*(2), 252–269. https://doi.org/10.1016/j.jml.2008.11.003

Swingley, D. (2009b). Contributions of infant word learning to language development. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3617–3632. https://doi.org/10.1098/rstb.2009.0107

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*(2), 147–166.

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*(5), 480–484.

Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition*, *71*(2), 73–108.

Taft, M., Xu, J., & Li, S. (2017). Letter coding in visual word recognition: The impact of embedded words. *Journal of Memory and Language*, *92*, 14–25. https://doi.org/10.1016/j.jml.2016.05.002

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.

Toro, J. M., Shukla, M., Nespor, M., & Endress, A. D. (2008). The quest for generalizations over consonants: Asymmetries between consonants and vowels are not the by-product of acoustic differences. *Attention, Perception, & Psychophysics*, *70*(8), 1515–1525.

Treiman, R., & Zukowski, A. (1996). Children's sensitivity to syllables, onsets, rimes, and phonemes. *Journal of Experimental Child Psychology*, *61*(3), 193–215. https://doi.org/10.1006/jecp.1996.0014

Van Ooijen, B. (1996). *Vowel mutability and lexical selection in English: Evidence from a word reconstruction task* (Vol. 24). https://doi.org/10.3758/BF03201084 van Ooijen, B., Bertoncini, J., Sansavini, A., & Mehler, J. (1997). Do weak syllables count for newborns? *The Journal of the Acoustical Society of America*, *102*(6), 3735–3741. https://doi.org/10.1121/1.420136

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review*, *97*(4), 488–522. https://doi.org/10.1037//0033-295X.97.4.488

Vihman, M. M., & Boysson-Bardies, B. de. (1994). The Nature and Origins of Ambient Language Influence on Infant Vocal Production and Early Words. *Phonetica*, *51*(1–3), 159–169. https://doi.org/10.1159/000261967

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic Phonotactics and Neighborhood Activation in Spoken Word Recognition* 1,* 2,* 3. *Journal of Memory and Language*, *40*(3), 374–408.

Von Holzen, K., Nishibayashi, L.-L., & Nazzi, T. (2018). Consonant and Vowel Processing in Word Form Segmentation: An Infant ERP Study. *Brain Sciences*, *8*(2), 24. https://doi.org/10.3390/brainsci8020024

Weber, A., & Scharenborg, O. (2012). Models of spoken-word recognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(3), 387–401. https://doi.org/10.1002/wcs.1178

Werker, J. F., & Curtin, S. (2005). PRIMIR: A Developmental Framework of Infant Speech Processing. *Language Learning and Development*, *1*(2), 197–234.

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, *59*(1), 114–132. https://doi.org/10.1016/j.jml.2008.03.001

White, K. S., Morgan, J. L., & Wier, L. M. (2005). When is a dar a car? Effects of mispronunciation and referential context on sound-meaning mappings. *Proceedings of the 29th Annual Boston University Conference on Language Development*, 651–662. http://titan.cog.brown.edu/research/infant_lab/White%20et%20al%20BU%202004.pdf

White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, *35*(4), 501–522. https://doi.org/10.1016/j.wocn.2007.02.003

Winskel, H., & Perea, M. (2013). Consonant/vowel asymmetries in letter position coding during normal reading: Evidence from parafoveal previews in Thai. *Journal of Cognitive Psychology*, *25*(1), 119–130. https://doi.org/10.1080/20445911.2012.753077

Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual Differences in Visual Word Recognition: Insights from the English Lexicon Project. *Journal of Experimental Psychology. Human Perception and Performance*, *38*(1), 53–79. https://doi.org/10.1037/a0024177

Zeguers, M. H. T., Snellings, P., Huizenga, H. M., & van der Molen, M. W. (2014). Time course analyses of orthographic and phonological priming effects during word recognition in a transparent orthography. *The Quarterly Journal of Experimental Psychology*, *67*(10), 1925–1943. https://doi.org/10.1080/17470218.2013.879192

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, *32*(1), 25–64.