universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

“Phono-stylistic variation in British pop songs: live performance vs. studio recordings”

Verfasserin

Barbara Bucher

angestrebter akademischer Grad

Magistra der Philosophie (Mag. phil.)

Wien, 2009

Studienkennzahl lt. Studienblatt: A 343
Studienrichtung lt. Studienblatt: Anglistik und Amerikanistik
Betreuer: Univ.-Prof. Mag. Dr. Nikolaus Ritt

Table of contents

1. Introduction

The fields of phonetic and phonological study are mostly restricted to spoken language. While this seems pretty obvious, it is remarkable how little research has been carried out on sung language. It is in so far remarkable as popular culture is filled with song and people are constantly exposed to sung language. The aim of this work is to show the phono-stylistic variation in pop songs, and for this matter five pop songs, all recorded in 2007 or 2008, have been chosen to be analysed. For each of these five songs two versions, one recorded in the studio and one performed live, will be considered.

The primary question of this work is whether the pronunciation of a singer differs in a live setting form that in a studio setting, and if so, how this variation can be explained. For this matter, Natural Phonology has been chosen as the theoretical framework for analysis, and predictions about the phono-stylistic peculiarities in the recordings will be drawn thereof. Natural Phonology is not the only theory to be considered in this work, and it will be shown where the theory meets its limits in terms of accounting for the phonetic variation in the songs.

This paper will begin with a short summary of the research carried out on the intersection between sociolinguistics and popular music, and it will be questioned as to whether the approach on the topic in the previous research is useful or valid. Then the theory of Natural Phonology, as first proposed by David Stampe and further developed by Wolfgang U. Dressler and Sylvia Moosmüller, will be introduced, and the theory's suitability to analyse the language of popular music will be argued. In addition, the field of Emotional Phonetics will also be focused upon shortly and its significance to pronunciation in pop songs will be shown.

2. Sociolinguistics and popular music

The rather narrow field at the intersection of sociolinguistics and popular music has only scarcely been explored. Even though the area offers a great range of data to be analysed, the number of papers published on this topic can be counted on one hand. One of the first papers that comes up when looking for publications on the pronunciation in pop songs is Peter Trudgill's "Acts of Conflicting Identity: The Sociolinguistics of British Pop-song Pronunciation" (1983)[1]. The only other article I ever found on the topic is Carl Johan Carlsson's "The Way They Sing It: Englishness and Pronunciation in English Pop and Rock" (2001), which follows Trudgill's basic concepts.

Trudgill claims that British pop singers adapt a particular "Pop-song style" which is an accent that differs greatly from their native one. There are (supposedly) specific 'tendencies' or 'rules' (Trudgill's inverted commas) which are employed when singing pop songs (cf. Trudgill 1997: 251). These tendencies or rules are then listed and condensed to six features that make up this specific pop-song style. They include the use of non-prevocalic *r*, the realisation of intervocalic /t/ as [d̬] and /aː/ becoming /æ/ (cf. ibid.: 252)

Unsurprisingly, Trudgill arrives at the conclusion that "[T]here can be no doubt that singers are modifying their linguistic behaviour" (252). This rather obvious point is then explained by *accommodation theory* according to Giles and, also, Le Page. It is then argued that singers, because they wish to identify with a particular group, modify their pronunciation accordingly to imitate - and thus appeal to - this group. As British pop music of the 1960's was highly influenced by American music, the features that appear in the pronunciation modifications may be attributed to American English. Yet the

[1] Please note that the page numbers here refer to a 1997 reprint in an edited reader. There has been no change to the article itself, yet the page numbers correspond to the 1997 version.

imitation of the influence group proved to be less than perfect and also inconsistent. The fact that Rock'n 'Roll's roots lie within the African American music tradition was known to British bands of the 1960's, and so they attempted to imitate African American Vernacular English. Seeing as the bands had little knowledge of this particular variety, however, they ended up producing an accent which more resembled General American than their actual target. (cf. Trudgill 1997: 253 ff.)

From the data, mostly Beatles and Rolling Stones records from the 1960s, Trudgill concludes that the use of non-prevocalic /r/ dropped dramatically with both bands over the course of time. While the Beatles realised almost 50% of possible non-prevocalic /r/ in 1963, by the end of the decade the number had dropped to less than 5%. It is also interesting to notice that the pronunciation of the band began to sound more like their native Liverpudlian (cf. ibid.: 258-261).

Trudgill then discusses the punk movement, which started in 1976 (bands include The Clash, The Sex Pistol and The Buzzcocks), and allowed for different pronunciations to enter the world of pop song singing. As punk's target audience was the urban working class youth, singers, even though some were not of such origin, changed the language in their songs to low-status and non-standard forms. Even though non-prevocalic /r/ was generally not realised at all, some features of American English are preserved in their singing style and American motivations conflict British ones.

Almost 20 years later, Carlsson picks up the topic and - while imitating Trudgill's approach - renews the data and analyses songs from the 1990s. To begin with, he also credits singing to be "considered in its own right" and each musical style to have "its own characteristics [...] in the way it is sung" (Carlsson 2001: 161). Carlsson then discusses the genre of Britpop /

alternative rock and looks at 24 songs. Unlike Trudgill, who observed six, Carlsson limits his study to merely four main variables: the loss of non-prevocalic /r/, the realisation of linking /r/, the pronunciation of intervocalic /t/ and the [ɑː], [æ] and [a] distinction in words like *after* (cf. ibid.: 162).

Carlsson's results confirm Trudgill's hypotheses of an ongoing change in British pop singing, as he found that the pronunciation overall proved to be more British than it was found twenty years earlier by Trudgill. Carlsson's results show that non-prevocalic /r/ was only used by a minority of 20% of the singers, linking /r/ was employed extensively, intervocalic /t/ was either realised consistently as [t] or [d] (never [ʔ]) and *answer* was realised with an initial [ɑː]. While he occasionally found other, regional features, such as northern [ʊ] in *but,* singers are neither consistent as individuals nor as a group (cf. ibid.: 164-166).

The conclusion at which Carlsson arrives is that the pronunciation in pop songs "can be seen as an attribute to the actual art form rather than a regional accent" (ibid.: 167). In a way, this particular pronunciation (yes, again a unique style of "pop-song pronunciation" is assumed) forms part of the musical culture, and is thus inherent to it.

While both papers show some very interesting points about the pronunciation in pop music, their attempt shows some flaws. Trudgill, more than Carlsson, tries to explain WHY the singers vary their pronunciation so much in sociolinguistic terms. Even though it can be safely said that the sung language of an artist certainly differs from the one that is spoken, Trudgill's comparison of what a specific singer does when singing to what a dialect group does when speaking is not quite satisfactory. It may be assumed that a singer, originating from a distinct dialect area, speaks with the significant

features attributed to his variety, yet Trudgill makes no remark about actually having heard the singers talk. Thus he assumes that when singing they differ from the way they speak, while he in fact has no evidence for this.

One example to illustrate how Trudgill's prediction somehow do not work can be found in a song by the London based band Dogs called "She's Got A Reason" in 2005. It includes a line, which is repeated several times: "I liked you better when you liked me aswell". The relevant word here is *better*, which according to Trudgill the intervocalic /t/ would be realised either as [t̬] or [ʔ] depending on whether they identify more with their musical role models or their audience. In the studio version of the song, this holds true and the /t/ is flapped in all but one instances, but when performed live this is a truly different matter, as it is realised as [tʰ].

Generally it should be stated that popular music in the UK has changed dramatically ever since the 1960s. The influence of American music is not as dominant, rather the British music scene has established its own standards and specifics. British pop music has been very successful, especially in this decade, and with British music being valued highly, Britishness itself - including pronunciation, of course - is more present in popular music. Singers do not feel the need to adapt an American accent because their musical role models are British.

The setting of the recordings is never mentioned, which is remarkable as Trudgill is usually so persistent on this matter and to him appears to be of no importance. As mentioned above, the aim of this paper is to show that the setting of a recording affects the pronunciation of pop songs and that a song recorded in the studio over the course of several days, even weeks, diverges vastly from the spontaneity of a live performance.

Furthermore, due to the fact that Trudgill and Carlsson claimed a "pop-song style" to exist, the songs were almost exclusively analysed in terms of few features and thus other interesting processes were ignored. There is only one instance where Trudgill goes beyond sociolinguistics and dialectology when he mentions that not all non-prevocalic /r/s are realised because they may disturb the "flow of the song" and "some phonological environments cause more difficulty than others. Most difficult, apparently, is the insertion of non-prevocalic /r/ in an unstressed syllable before a following consonant" (Trudgill 1997: 257). This discrepancy is the starting point of this paper, which will try to explain the motivations for phono-stylistic variation in British pop music.

3. Natural Phonology

3.1. Definitions of Natural Phonology

Natural Phonology, henceforth also NP, is a theory on phonology which is formed upon actual speech. One of its key elements is "the empirical centrality of external evidence" (Gibbon 2007: 83). NP includes speech data from various speech groups in various settings and goes beyond the (assumed) *ideal speaker* (cf. Foltin & Dressler 1997: 5). And according to Donegan & Stampe "in the case of natural phonology this means everything that language owes to the fact that it is *spoken*" (1979: 128).

Foltin and Dressler offer a very short, yet precise definition of what NP does:

> Die Natürliche Phonologie (NPH) sucht die Erklärung für [...]
> Erscheinungen in Bereichen, die außerhalb der Phonologie,
> [in diesem Fall] der Phonetik liegen[...] (Foltin & Dressler 1997: 5)

Unlike other phonological theories, NP claims to be a *natural* theory and it shows language as "a natural reflection of the needs, capacities, and world of

its users[...]" (Donegan & Stampe 1979: 127). Thus language depends on its speakers and hearers, and NP accounts for this by dealing "with the production and perception of speech sounds - as well as with their alternation and variation [...] (Donegan & Stampe 2008).

In speech a speaker does not necessarily hear what she or he actually says but hears the sounds that were intended (cf. Donegan & Stampe 1979: 159). Natural Phonologists call upon Baudeouin de Courtenay and his definition of a phoneme as a sound intention (cf. Donegan 1978, Donegan & Stampe 1979, Stampe 1984, Moosmüller 2007 etc.) when they relate a sound to its underlying representation. Due to speakers' physical limitations, to arrive at the actual output (no matter how different this may be from the speaker's intention) *processes* are applied. Vowel nasalisation, for instance, is never intended[2] in English, yet it occurs naturally in vowels that precede nasals. Intention here means that a speaker aims for an articulatory target and arrives at an output which has been affected by processes. So even if the nasal itself is lost, and it may not even be realised at all, the vowel is still nasalised despite the lack of a nasal to follow. This, for instance happens in one of the songs analysed later, "Jacqueline" by The Coral, where *down* is realised as [dãː] with the nasal being obviously absent but the vowel remaining nasalised. The singer's articulatory target involves a final nasal, in terms to realise the sound the preceding vowel is nasalised, and despite the nasal not being realised it can still be heard in the quality of the vowel.

As Peter Ladefoged put it in 1992:

> [...] it would be possible to devise abstract phonologies that [describe] all sorts of pretty patterns that might have no physical correlate of any kind (Ladefoged 1992: 165)

[2] This assumes "normal" language use and does not take other, language conscious instances into account, like the imitation of a French accent etc..

While this might hold true for some phonologies, the aim of NP clearly lies elsewhere, as one aim of phonological theory is, to quote Ladefoged once more, "to help to explain why languages have the sounds they do." (Ladefoged 1992: 166).

The very basic claim of NP is its universality. In that sense phonological processes and rules are included in all languages, yet they are language specific. That means that one process may be suppressed in one language yet not in another. For example, the process of final obstruent devoicing is oppressed in English but not in German. (e.g. compare German *Land* /lant/ and English *land* /lænd/.

The fact that NP draws its conclusions from actual speech data is vital for this thesis. It seems that NP, by creating a phonological theory based on what happens when speaking and listening, offers a perfectly fitting framework to observe the pronunciation in pop songs. Because of the fact that it is purely descriptive and not prescriptive, the theory will allow for all phonetic peculiarities to be included.

3.2. How Natural Phonology works

3.2.1. Processes and rules in Natural Phonology

In 1969, David Stampe first introduced the concept of *processes* to phonology, about most of which he drew his conclusions from child language. The purpose of these processes is to "systematically but subconciously adapt our phonological intentions to our phonetic capacities[...]" (Donegan & Stampe 1979: 126). In that sense, languages adjust according to our vocal tract and what is anatomically possible, thus processes are a response to phonetic difficulties (ibid.: 136). Furthermore, "[p]honological processes serve the

communicative function of language" (Dressler 1984: 31) by making it easier to be pronounced and perceived.

According to Stampe, language is "the residue of an innate system of phonological processes, revised in some way by linguistic experience." (Stampe 1969: 443) These processes come in contradictory sets and in some cases when they may overlap, either process may be suppressed or partly suppressed (ibid.: 443f.). It is in this sense that children by learning their first language "systematically inhibit the process types which are (or become by maturation) available to them according to the inhibitions of the linguistic norms they are exposed to" (Dressler 1984: 30). This in short means that learning in NP can be described as the mastering of inputs of natural processes which the words of the learner language require (cf. Donegan & Stampe 1979: 140).

A child, which Stampe assumes to be in a linguistic unbiased state[3], may apply all processes and, step by step, will learn to suppress specific processes and arrive at a stage in which its speech will resemble adult speech. If a child fails to suppress an innate process applied in standard language, phonetic change occurs (ibid.: 448). Dressler describes a "typical scenario of diachronic change" (Dressler 1984: 34) in the following way: first an assimilatory process occurs in casual speech, then it enters formal speech and finally becomes obligatory.

Processes are unordered and all occur on the same level, at the same time. This contrasts them strongly to phonological rules, which have no connection to articulation and are thus not naturally motivated but "function to assure

[3] Unbiased here means that the child has not yet learned which processes apply in a language.

the necessary grammatical complexities we associate with allophony and allomorphy" (Bjarkman 1986: 81).

Another important aspect of processes is that they can be divided into two groups, the one being *prelexical* and the other *postlexical* (cf. Dressler 1984, Dressler & Moosmüller 1991). While prelexical processes determine the phoneme inventory, i.e. phonotactics, and are obligatory, postlexical processes are either obligatory or phono-stylistic. Obligatory postlexical processes can be further subdivided into being phonemic (e.g. final obstruent devoicing) or allophonic (e.g. context-sensitive aspiration). Phono-stylistic processes, which may be optional, characterise, for instance, casual or fast speech and are applied in such setting (cf. Dressler & Moosmüller 1991: 137). As the title of this paper already suggests, phono-stylistic processes will be most important in the variation that occurs in pop songs.

Processes only alter one phonetic property, or feature at a time to overcome a difficulty and try to be perceptually similar to the original sound. Contrastingly, rules can substitute a phoneme by a completely different one, which is possible because rules are not a response to phonetic difficulties and are learned. Some rules of the English language are, for instance, velar softening or tri-syllabic laxing (or shortening), and these rules are applied because of convention, and even though through habitual use their application may be subconscious, rules were once conscious processes which have been conventionalised. Unlike processes, rules are always obligatory and because their application does not change the meaning of a word, they do not depend on a specific style. Furthermore, rules are applied before processes. (cf. Donegan & Stampe 1979: 137-145, 156). Rules are also (mostly) general, transparent and recoverable.

As for one reason why rules are hardly mentioned in NP, Dressler assesses that the boundary between rules and processes is somewhat blurred and not always easily drawn (cf. Dressler 1984: 30). In another paper, Dressler and Moosmüller even claim that rules are excluded from NP because they have a "morphological rather than a phonetic function" (Dressler & Moosmüller 1991: 136). It lies within the nature of rules that they will not be considered here. Their application is pre-lexical and always obligatory, which is why they should be applied in either recording of the pop songs. Variation in terms of rules is not to be expected.

3.2.2. Fortitions and lenitions

Processes may be categorised into three types: *prosodic processes*, which map phrases and sentences onto prosodic structures, *fortition processes*, which strengthen an individual segment and *lenition processes*, which make segments easier to pronounce by making them "less distant" to the surrounding segments (cf. Donegan & Stampe 1979: 142 f.). As the prosodic mapping in "verse or music are special cases" (ibid.: 142), only the latter two types of processes will be relevant for this work. The language in songs differs prosodically from spoken language in a way that relevance would be questionable. Stress patterns in a songs are determined by a songs rhythm and not by the rhythm of speech. Intonation is determined by melody and cannot be compared to intonation patterns in spoken language. This is why, here, the focus will be more on segmental processes, rather than on suprasegmental ones.

Donegan's work from 1979 concentrates on processes in NP and she tries to explain the need for fortitions and lenitions in language in the following:

> On the phonological level, the conflict is between the need to maximize the articulatory and acoustic properties of individual

> sounds -- mainly in the interest of perceptibility -- and the need to minimize the articulatory effort required to produce these sequences of segments that form syllables, words, and other units. (Donegan 1979: 20)

These two different requirements may result in opposite changes, as the needs of the speaker and the listener may be motivated conversely. While both kinds of processes - lenition and fortition - are equally common in speech, Donegan and Stampe argue that certain types of processes are linked with particular settings. The most prominent claim is that lenitions occur in fast or casual speech, whereas fortitions and hyperarticulation are commonly attributed to careful and formal speech (Donegan 1978, Donegan & Stampe 1979). Predictions about the application of processes in pop songs cannot be made from this claim, as will be explained in the following.

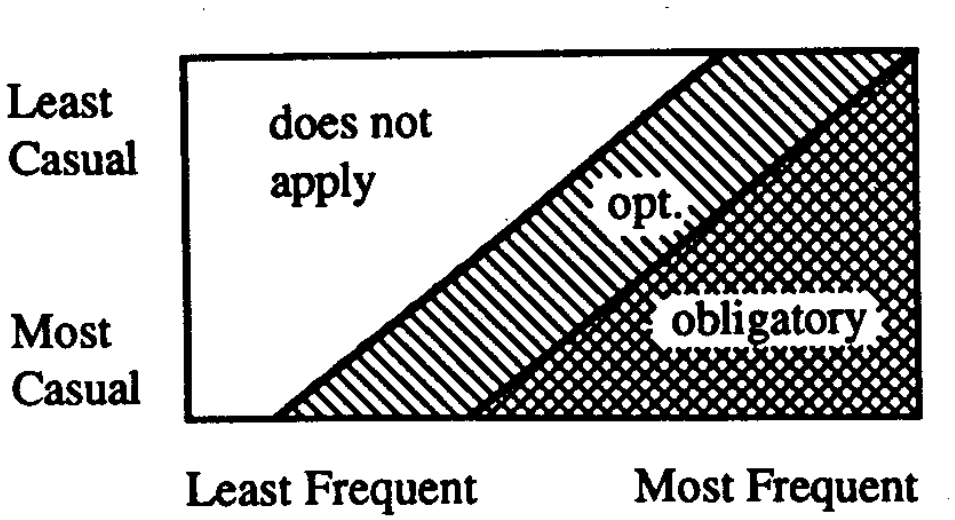
Fortition processes, also referred to as strengthening, increase some phonetic properties of a segment and thus are "often increasing the contrast between the segment and its environment" (Donegan 1979: 21). Relating this process to the pronunciation in pop songs, it can be expected that fortition processes occur frequently in a studio recording, because this is the recording that an audience listens to repeatedly, on the radio or on a CD. Fortition processes can make a recording be understood more easily, which is also relevant in a live setting, where there is background noise.

Processes that make segments more pronounceable are called *lenitions* and ease articulation by "assimilating the properties of one segment to those of a neighboring segment, by deleting segments, and by substituting segments rather than sequences" (ibid.: 21). When lenition processes apply, the contrast between a segment and its environment is weakened or eliminated. Lenition processes are less likely to be applied in studio recordings because they make lyrics less perceivable. In a live recording, where a singer is probably more focused on performing than on pronouncing words clearly, lenition processes

are expected to occur more frequently. There is also the fact that a live audience usually knows the songs that are performed, and thus lenition processes on behalf of the singer do not obstruct understanding.

Rhodes claims that a single word can have a wide variety of "pronunciations based on how casually it is pronounced" (Rhodes 1996, 242) and also the frequency with which words are used affect weakening processes that may be applied (cf. Fenk-Oczlon below, following section). So in a song where a word is repeated over and over, variation is to be expected.

(1) The applicability of a natural process



(Rhodes 1996: 245)

From this figure (1) one can see the gradual application of a phonological process according to the frequency of the word and the level of casualness of the setting[4]. Whether the setting in a studio is more formal than in front of an audience is debatable (the term *casual* seems inappropriate in this context) but the frequency of words also affects the pronunciation in pop songs.

[4] The applicability of a process is, however, not always influenced by the word's frequency, and while the image shows one possibility, it by no means explains the application of all processes, as for example, German final devoicing which occurs in least frequent words in least casual setting. An instance where the process, according to the image, should not apply.

Words that are very frequent in speech in general, such as *the*, are prone to reduction, but also segments that are repeated several times in a short interval are affected. In the song "Ice Age" by Good Shoes, for instance, almost every line of the two verses begins with *I'm* or *I*, summing up to a total of 19 times. The majority of these segments is reduced to [ə] in both the live and the studio recording. Frequency may be a factor that influences this,

3.3. Natural Phonology and semiotics

From the late 1970s onwards, Wolfgang Dressler began to introduce the Peircian notion of semiotic *figure* and *ground* to phonology. From the semiotic concept he was able to derive the notion of processes which he categorised as *backgrounding* and *foregrounding*, which in the Stampean sense would be classified as lenitions and fortitions. Instead of placing the motivation of these processes in the phonetic qualities of speech alone, in Dressler's theory the processes of backgrounding and foregrounding make specific parts of speech stand out of the ground or blend in. This is similar to how Donegan and Stampe propose fortition and lenition to work, but this is expressed more clearly with semiotic terms. Also the metatheory of semiotics is aimed to help to put NP into a larger context, and if language is assumed to be a system of verbal signs, then this should also work for phonology (cf. Dressler 1984: 32).

Applying this to phonological processes, it follows that the input, i.e. the phoneme, is the *signatum* (signified) and the output, i.e. allophone, is the *signans* (signifier). Ideally, allophones should be distinguishable from each other, which mainly holds true for formal speech but fusion and deletion processes are common for casual speech. Furthermore, the relationship between signified and signifier should be transparent and reliable (cf. ibid.:

35f.). It would be (semiotically) confusing if there were a process which lead to /l/ being realised as [f], as the relationship between signified and signifier would be non-iconic and thus no longer functions. Another constraint on perception, and production as well, is the suppression of an obligatory process.

A detailed account of how the inclusion of semiotics enhanced the NP theory, Fenk-Oczlon strove for an extensive list of foregrounding and backgrounding processes, summarised here in the following table (2) (cf. Fenk-Ozclon 1989).

(2)

Backgrounding	Foregrounding
assimilation	dissimilation
vowel reduction	diphthongisation
deletion	insertion
in final or mid position	in initial position

This is not merely a renaming of the same concepts, but adds more to the scope of the processes. Foregrounding includes more than just fortition, as well as backgrounding goes beyond lenitions. An example drawn from popular music would be, for instance, Ocean Colour Scene's "If I Gave You My Heart" where the words *stood on* are realised as [stʊ ʔɒ̃n]. To arrive at this output, the easiest explanation would be to say that the final /d/ has been substituted by a glottal stop. When listening closely, however, one can hear a brief break just prior to the glottal stop, yet none after. It is because of this that the glottal stop cannot be a replacement for /d/.

Instead the final /d/ of *stood* is elided[5], clearly a lenition process, then a glottal stop is placed before *on*, which is a common process in English referred to as *hard attack*, clearly a fortition process. Both processes combined result in the foregrounding of the segment and the effect simply seems to make the segment stand out instead of either easing production or perception. If sung as [stʊd ɔ̃n] the sounds are neither ambiguous nor "difficult" to produce (the voiced stop is in intervocalic position and is not naturally devoiced because of its neighbouring sounds) yet the singer still opted for a foregrounded realisation. This is a foregrounded realisation, because of the larger context of the song. The song's time signature is 6/8, and each syllable on the beat is stressed and followed by five unstressed ones. In this particular segment, the beat is on *stood* and because there is a hard attack in *on* there is a brief pause and a sharper contrast between the two words, which underlines the rhythmical pattern and at the same time makes the words more perceivable.

3.4. Natural Phonology and sociolinguistics

The incorporation of sociolinguistics into NP dates back more than 25 years and adds an interesting aspect to the theory. While it may not be initially obvious, NP and sociolinguistics are tied together quite closely, as the theory accounts for actual human speech and does not exclude the speaker from the equation. As humans are social beings, each individual is embedded in society, which therefore is a relevant factor.

Dressler & Wodak's 1982 study on Viennese German was conducted within the NP framework and analyses its data in terms of rules and processes. They collected data from various settings: formal and informal interviews,

[5] Even though some dialects allow for final voiced alveolar stops to be glottalised, in this segment, the glottal stop does clearly (as in audibly) not belong to the syllable of *stood*.

spontaneous interactions in court, patients in psychotherapeutic groups etc. The main aim of the analysis is to find how differences of formality in the speaker setting affect their speech. For this matter they first explain the different phonological rules of dialect and standard and then observe the application or suppression of processes in different settings. According to Dressler & Wodak's view "phonological styles represent sociolinguistic levels of formality" and thus it can be predicted when processes are applied or not (cf. Dressler & Wodak 1982: 339-351).

When speaking, a speaker may decide to use a different variety and by doing so applies a different articulatory target arriving at a different output. So, for instance, if a Tyrolean speaker decides to pronounce the second person singular of *be,* which is *bist* in German, the dialectal input would lead to the following output : [b̥ɪʃ]. If the speech situation required the speaker to speak Standard Austrian German, a standard input would be used and result in an output somewhat like [bɪst]. The two different outputs may be attributed to a specific speech situation or style, but the main point is that even though the dialectal output may appear in a more casual or colloquial setting, it is not derived from the articulatory target of the standard variety through the application of lenition or backgrounding processes but by using a different target in the first place.

This fact has to be considered when observing the data from the songs because there might be some peculiarities that would indicate a unique set of processes applied if derived from the standard target, but are fairly easily connected to a dialectal target. For example, singers from the North of England may not have a distinction of /ʌ/ and /ʊ/, and an instance when *but* is realised as [bʊt] cannot be explained by fortition or lenition of the standard variant but by a singer aiming for a different target. The choice as to

when which target is striven for may be conscious and can have a foregrounding or backgrounding effect.

More recently, in 2007 Sylvia Moosmüller stressed that Donegan & Stampe's "classical" attribution of processes to the sheer casualness or formality of speech was not sufficient and that processes are much more dependent on the social setting than supposed ease of articulation.

Especially with developments in experimental phonetics it has been proven that ease of articulation is something so specific to one speaker, actually to one vocal tract to be precise, that this cannot be generalised to a whole set of speakers. The economisation of speech as the one goal of lenition processes is not sufficient because each speaker has individual needs and the assumption that a process which eases articulation for one speaker, is also true for another is not true (cf. Moosmüller 2007: 8). What truly happens, Moosmüller suggests, is that in a particular setting some processes are agreed upon and hence applied:

> Therefore, in a casual speech situation, a person speaks the way it is expected from her or the way she thinks it is expected from her, but not according to a principle of least effort. (Moosmüller 2007: 4)

As mentioned above, classical NP explains fortitions to enhance understanding in terms of the listener and lenitions ease articulation for the speaker. Moosmüller, however, questions the applicability of a dichotomous concept of speaker and listener whose needs are supposedly antagonistic.

A speaker can only influence the application of processes to a very limited extent. The interactional situation determines the *figure - ground* contrast and the speaker adheres to that as good as she can. An individual speaker's

interest lies not within speaking economically but the focus lies on the success of communication. Thus the processes which are applied are "motivated by the social group not by the individual's needs to ease production" (ibid.: 8).

4. Emotional phonetics

While the theory of Natural Phonology can be used to predict phono-stylistic variation in pop songs, another topic in phonetics should not be overlooked, and will be introduced briefly. Emotional Phonetics is an area of study that investigates how or whether emotions are encoded in the acoustic signal. As the singers of pop songs can be expected to be emotionally involved in their performance of some sort, may it be anxiousness or anger, these emotional cues may affect their pronunciation. Emotional involvement is dependent on the singer, and more involvement is to be expected when a song carries a strong message. If a punk band, for instance, wrote an aggressive song about the mischievous demands of capitalism, the aggression will be heard in the singer's voice.

In research on spoken language there are claims that "different emotions can be recognized on the basis of vocal cues alone" (Banse & Scherer 1996: 614) and this means that a listener can determine the emotional state of a speaker not only by what is said, but also by how the signal actually sounds. There are emotional states which greatly affect respiration and articulation, for instance the emotional state of stress[6], and this again influences the properties of the acoustic signal (cf. Scherer 1981b: 182). Parameters which are involved in signalling emotion are:

[6] Stress is difficult to define and generally is a state where a person is upset through psychological or other stimuli. Stress can be caused by mental exercise, drugs or danger. For a more detailed discussion on the problem of defining stress see Scherer 1981b: 171f..

(a) level, range and contour of the fundamental frequency

(b) the vocal energy/amplitude

(c) distribution of the energy in the frequency spectrum

(d) the location of the formants

(e) tempo, pausing etc.

(cf. Banse & Scherer 1996: 615f.).

In studies on the effect of emotions on phonetics different methods of gathering data have been applied. There are studies which asked actors to perform emotions, something which could be compared to the live performance of a song, sometimes people were asked to tell stories which set them into a certain emotional state (cf. Scherer 1981a: 201-203). Recognition of the emotional state was tested with a set of people listening to recordings of speakers in specific emotional states, yet this only supplied mixed results. When asked to define the speaker's emotional state, the accuracy of judgement is in some studies as high as 80% and in others less than 30%. (cf. ibid.: 208-210). Spectrographic analyses for emotions like pride or cold anger did not provide satisfying results either. However, for some extreme and contrasting states, like anger, excitement or sadness, results appear to be conclusive.

The methodology which uses actors as performers of emotions has been criticised, as these emotions were then not "natural". Yet Banse and Scherer argue that "real-life" emotions are just as natural or unnatural as staged ones. According to them, one cannot tell a difference between a performed emotion by an actor or an emotion a close friend expresses in private (Banse & Scherer 1996: 618). What is interesting is that the distance to Natural Phonology seems not that far after all,

> "[a]s the nature of the verbal utterance serving as carrier of the vocal emotion expression greatly influences the acoustic parameters (the vocal channel being jointly used for linguistic and emotional signaling)" (Banse & Scherer 1996: 618)

If emotion is naturally attached to the speech signal then it would seem safe to assume that natural processes are affected by emotion. Unfortunately the parameters for recognising emotional cues in the speech signal are more suitable for vowels than for consonants and the distribution of formants is rather difficult to judge without technical assistance from a computer. There are, however, some parameters which may be relevant in the analysis of the songs.

Loudness is an aspect of speech that is not a process accounted for by NP (it serves in terms of foregrounding and may ease perception, this depends on the greatly on the setting, i.e. whether there is background noise, the listener is paying attention etc.) but is important here. Generally emotions like anger and joy are connected to increased energy. Because of the unique circumstances of recording a song, loudness of the vocals can only vary as much as the whole band varies. That means even though the passage may be sad, expressing this tristesse by singing more quietly would result in simply not being heard at all. The general effect in pop songs then is the decrease of the general volume. Only then can the voice be used more softly and thus express a sad tone. The opposite holds true for uplifting or angry songs, which are usually more upbeat as well. If anger is to be expressed it can be achieved in terms of loudness or simply more energy. In "Tell Me What It's Worth", the second verse expresses anger, for example the line "kill, kill, kill when everything starts to suck", is sung loudly and also with a lot of energy which is most noticeable in the strong aspiration of consonants.

Needless to say, some songs do not need phonetic cues to convey their emotional message, as there usually are lyrics attached which may just fulfil

that task but the music can override this[7]. So if the Scottish band Travis were trying to fulfil all acoustic criteria for sadness in their song "Happy", slowed it down and reduced it to an unplugged version, the chorus line "I'm so happy, 'cos your so happy" may very well be perceived as sad. Arrangement, volume and tempo of the song are essential on its emotional impact. The lyrics of a song are subordinate to this because they are not as obvious. To illustrate this with another example, one where a contrast between lyrical content and music was created intentionally by an artist, is The Jam's "Town Called Malice". There the upbeat and energetic musical frame distract the listener from the sinister lyrics, which (also due to recording quality and low level of the vocals) are not understood when listening only once or twice.

I have had numerous, seemingly endless discussions, with people who were either convinced that song lyrics are the most important part of a song or others who claimed not to pay attention to the words that were sung because they did not matter to them. The attention song lyrics attract is dependant on the song itself, so an upbeat, dance song, for instance, will not have as many attentive listeners in terms of their lyrical content. As The Pipettes put it very precisely in their song "Pull Shapes" : "I just want to move, I don't care what the songs about".

On the other hand a song does not even need words to submit a general feeling of happiness or joy, nonlinguistic vocalisations work just as fine, for instance in Starsailor's "Keep Us Together" the chorus is comprised merely of "oh oh oh oh", repeated thrice. The emotion lies within the melody and the acoustic properties of the *oh*, not within the linguistic message. These example nonetheless show the many layers that lie within the vocals of a pop song. Emotions are one of them and emotional phonetics may not be left out

[7] Music was also used in the mood induction technique to put speakers who are supposed to speak emotionally loaded in a specific frame of mind (cf. Scherer 1981: 203f.).

here as a contribution factor, yet it is only one piece of a three-dimensional puzzle.

While in the study there will only be few references to Emotional Phonetics, it is nonetheless very important to keep it in the back of one's mind when analysing the data. The lack of references is not due to any invalidity or inadequacy of the theory, but rather ensued from the songs being observed out of context. This made it difficult, if not impossible, to correctly observe and judge emotions as they were performed by the singers, which only allowed for tentative speculation.

5. Methodology

The basic aim of this work is to apply linguistic theory on phonetics and phonology on the five pop songs[8] chosen to be analysed here in order to account for phono-stylistic variation. While the set of data is small, it should provide sufficient insight into what happens phono-stylistically when songs are recorded in the studio and when they are performed in front of a live audience.

In order to work on the phonology and phonetics of the songs they had to be transcribed. Because even slight variation is vital here, the transcription is narrow and generally follows the symbols of the IPA with their according usage. The transcription will be presented in the following way: one line in black print which is simply graphic representation of the lyrics, underneath is a line in pink, which is the transcription of the studio version, and

[8] The five songs were played on a computer and listened to through in-ear headphones in either MP3 or AAC format, with a standard sample rate of 44,100 kHz and a bit rate ranging between 128 kbps and 256 kbps. Even though the songs were recorded using a higher bit rate, these rates are common for legal music downloads and do not affect the sound quality to an extent where it would distort perception in any kind of way, rather this enhances the downloadability and keeps required storage space to a minimum.

underneath this is the transcription of the live version in light green. For convenience, the lines have been split up into several blocks, usually at logical, rhythmical breaks in the song. That way, the transcription, which runs at different typographical lengths, catches up and can be followed more easily. These transcriptions can be found at the end of this paper, in the appendix.

The transcription, which is purely impressionistic and relies on the ear of the observer, will be analysed in section 6. The main focus of the analysis will be on the phonetic differences between the two recordings caused by the application or suppression of fortition and lenition processes as discussed above. Other aspects of pronunciation, which cannot be attributed to a process will not be left out, such as different realisations caused by different articulatory targets.

While specific phonetic features are interesting, in order to achieve an overview of how often processes are applied either live or in the studio, the transcriptions have been analysed in terms of seven specific processes, the applications of these processes have been counted and transposed into a table. The seven processes are, subordinate to foregrounding: *devoicing*, *aspiration, diphthongisation, insertion;* and subordinate to backgrounding: *voicing, monophthongisation* and *elision*. These processes were chosen because they occur in all the songs and comparisons can within the songs as well as between the songs can be drawn. For each song one table including the seven processes was created, giving numbers of how frequent the processes occur live or in the studio; additionally, the discrepancy between those numbers has been calculated. All those numbers have then been combined to an additional table giving the total numbers of all five songs.

There are, however, other processes, which will be discussed, including *affrication, lowering* and *lengthening*. Seeing as those processes do not occur as frequently as the other ones, and some are restricted to one song only, their application has not been counted and summarised in a table. A mere verbal discussion and analysis will be sufficient.

The point of the analysis and discussion is to answer the question whether there is phono-stylistic variation in pop songs when they are performed in the studio and when they are performed live. The ordering of the variation under the general terms foregrounding and backgrounding allows for conclusions to be drawn more easily than if there was no ordering. It will be shown whether the predictions which have been made in section 3. on Natural Phonology (foregrounding processes will be prominent in studio recordings, while backgrounding processes will be applied more frequently in live situations) will hold true or whether the results draw an altogether different picture.

6. About the recordings

6.1. Why the recordings were chosen

It proves not to be easy to find suitable data for linguistic analysis in popular music, and while the actual quantity of high quality recordings is vast, some parameters have been applied to pin the amount down to a very few songs. As the number of recorded pop songs, live and in the studio, is so great, one could have easily picked from a heavy back-catalogue of at least 50 years.

Two main principles were important when choosing the songs. When the songs were recorded and where they were recorded appeared to be the most significant factors. Time is an important issue, as recordings from today are much better in their quality, and quality here is the essential factor. Digital recording and mixing techniques allow very little deviation of the actual sound and thus makes it clearer and individual voices on the recordings are more distinguished. This holds true not only for the studio recordings but also for the live recordings. For this reason all recordings observed here date from 2006 to 2008, a very recent and narrow time span.

It is relevant what time span lies between the two different recordings of the song which is to say that how much time between having recorded a song in the studio and a given live performance have passed. Or vice versa, which seems to be hardly ever the case. For this work, songs have been chosen that were released as a studio recording merely a few months before the live recording. This is important because singers sometimes change their singing accent over time. In an analysis where the main changing factor is the setting, and which effects this has on pronunciation, and additional factor of variation over time would complicate the matter.

One of the artists whose recordings are of interest here is Paul Weller. Having been a recording artist since 1977, his work is as broad as it varies. It is quite easy to hear a change in his accent from the late Seventies where he was in The Jam, a (supposedly) punk band, which turned more pop towards the Eighties, towards the more jazzy The Style Council in the mid-Eighties and then further accent changes are audible throughout his solo career from the beginning of the Nineties. When one listens to recordings of a The Jam song that were made in the Nineties, the accent differs very much from the Seventies recording. Such a comparison is interesting as such, but not useful here, as said before, the setting is the focus of the analysis, not time. The aim

with the sorting process was to find songs that were performed live very shortly after they had been recorded in the studio.

The issue of where the recordings have taken place is more important with the live versions than with the studio ones. The quality of live recordings is sometimes so poor one can hardly hear the voice at all. The studio recordings are all of the same high quality, they serve common recording standards with a bitrate of 128 kbps or higher, and for this purpose, in terms of quality, it can therefore be neglected in which studio they were recorded. Yet the live performances were all recorded in the same venue, the London *Koko* club, which is of some relevance here.

All the songs having been recorded in the same venue ensures that all artists were able to use the same high quality equipment and all faced a same sized audience limited to 350 people, even though Koko is a venue that holds up to 1,410[9]. 300 of the 350 tickets were given away in competitions[10] and none of the remaining tickets were sold. For all of the artists that are surveyed here, this is a fairly small venue. All live recordings were made within the framework of the iTunes Festival London, which was a month long festival in 2007 and again in 2008 with many bands performing at Koko. After a short period of time the recordings were available for purchase as a download on the Apple iTunes Store.

The five songs are:

2007	"Ice Age"	Good Shoes
	"Jacqueline"	The Coral
	"Sexy In Latin"	Little Man Tate

[9] http://www.koko.uk.com/files/prodbible.pdf (03.01.2009)

[10] http://www.macworld.co.uk/ipod-itunes/news/index.cfm?newsid=18270 (03.01.2009)

2008	"All I Wanna Do (Is Be With You)"	Paul Weller
	"Tell Me What It's Worth"	Lightspeed Champion

6.2. The recording process

6.2.1. Recording in the studio

As the setting hypothesised to be a key element in the variation in pop song pronunciation, it is therefore necessary to elaborate on the recording process itself, and to describe the different settings in a studio compared to a live venue. There are many ways to record, yet there are a few factors which hold true for almost every studio and live setting, and it is those which will be considered here. The recording process in a studio can differ vastly from a live performance, but it does not have to.

Generally, there are different ways record a song. Either one records "live", which here means all instruments and vocals perform at the same time and are recorded over various microphones and then mixed. This very much represents an actual live performance. There can be alternations to this, which means that there may be more than one take, i.e. recordings, or overdubs, which are re-recordings of a single track. This is a very common practice. First the band records a song together and later on little things are added or are altered.

More rarely with indie[11] and rock music, one track is recorded after the other and only then layered to form the complete song. This allows many takes and alternations as one track is clearly assigned to one voice. Recordings of this type often include many programmed elements and effects that are used

[11] Indie is a term applied for music which is guitar based and was originally promoted by independent labels. With the success of this type of music the term was used more broadly to refer to the specific style rather than the band's label policy.

with the help of a computer. Also, this kind of recording can be done alone without the help of a producer or an engineer and is often chosen as a home recording amateur method.

As mentioned above, the "live" recording process is often preferred by indie artists because it also creates an atmosphere of intimacy and authenticity. As Moore put it in 2002 with reference to Paul Weller: "There is his practice of recording "live" in the studio, i.e. with an absolute minimum of the overdubs, multi-tracking and other devices which "cheat" the listening ear" (Moore 2002, 212). The alleged superiority of live performance is also touched upon in Bannisters 2006 article where he discusses how the indie scene emphasises "liveness" and minimalism as much as amateurism in their recordings (Bannister 2006, 83f.).

This emphasis on a particular sound, meaning to sound cheap, with fuzzy background noises and distorted vocals, is surely a reference to analogue recording times, yet nowadays not as easily achieved. While some bands choose old equipment, instruments, amplifiers, effects, mixers etc. to master an "authentic" Sixties or Seventies feel, others use built in computer programme effects. Using these computer programmes is cheaper than buying "vintage gear", which is difficult to find and additionally so sought after that prices exceed most aspiring artists' imagination.

There are two programmes that are used in studios world wide, Logic Pro and Cubase. Both come in home user friendly Express or Essential versions that are limited, yet affordable. The full versions of these programmes offer recording possibilities beyond what could have been achieved with masses of hardware only fifteen years ago. The range of processes is so large it takes a professional to master and apply them in way required by professional recording artists.

This leads to one of the most important factors of studio recording: the presence of a producer and engineer. While some bands and artists prefer to produce their songs themselves, others hire a professional producer who gives ideas about how to record a song and what it could and should sound like. The engineer, on the other hand, is responsible for making the machines and electronics work. In a studio there is a room for recording, with various separated booths for drums or vocals, and a mixing room where the producer and engineer work. Those two rooms are almost always separated by a a glass wall which deadens sound.

It is important to notice that the producer often enjoys some status of authority, he or she is the person with the plan and the musicians are supposed to act accordingly. In that respect the hierarchy changes in a live performance where the musicians are up on a stage and the audience follows them. In the studio, the producer is also a mediator of the sound, whereas live there is no such mediation. This holds not completely true, of course, because in a live setting there is also a sound engineer who employs the mixing desk. It is her or him at whom some weird and outraged gestures of musicians are directed when they are unhappy with what they hear.

The influence of the producer on the performing artists must not be underestimated. If the producer is not happy with a result, it has to be recorded again. I vividly remember reading an interview with Danny McNamara, singer of the band Embrace, some years ago, in which he told how the producer almost brought him to tears when recording a song. Apparently the producer asked him to imagine his girlfriend dead to achieve the level of sadness in his voice. This, of course, ties in with the framework of Emotional Phonetics, as the producer directly influenced the singer's output by toying with his emotions. Hence, variation is not only influenced by the studio setting but also by a third party, here in form of a producer.

6.2.2. Performing live

As mentioned above, the setting of a live performance can hardly be compared to the setting in a recording studio. Firstly, the positioning of a performer on stage is unlike the one in the studio, and what is even more important, the performer faces an audience.

When performing live, the singer has to accomplish being heard over the noise of the crowd and the music the band is playing. This may at times be difficult, as the reaction of the listeners may distract the performer, or technical problems may interfere. While the interacting aspect of live singing can be named the cause of phono-stylistic variation and this variation may be initiated subconsciously, it should not be forgotten that a singer can also consciously adjust her or his speech, in order to achieve a comic effect or draw attention to a specific word or line.

The notion of variation is strong with some artists, and one of them, Paul Weller, whose song forms part of the data, puts emphasis on this on the LP sleeve of his 2008 collection of BBC recordings, where he wrote:

> As for a live show I try and vary it as much as I can from the last time, but you can't really prepare though, you don't know until you get on that stage what it's going to be like, what the sound's going to be like, how the crowd will react to the songs, whether you'll be any good or not, it's all up in the air but that's what makes it brilliant and makes you come back for more. That's what I do, I mean I write and make records, but essentially what I do is play music live to people.

In indie music, the live performance tradition is valued highly and almost demands that a live performances should exceed studio recordings. Judging from the price ranges of live tickets compared to those of recordings, this is understandable. Hardly any live performance sounds like a studio

performance, and it can be doubted whether this ever is the performer's aim anyway. Going to, and indeed playing, a concert which sounds exactly like the record would be boring. Of course this does not hold true for all performers and performances, yet variation to some extent is fairly common, as long as it does not impede any communicative function. Seen in a more narrow linguistic context, the fact of variation also influences phonetic outputs. This has been established in linguistic research, and, in the words of Jennifer Pardo wrote:

> Speech is variable in its realization, both within and between talkers, despite apparent consistency in perception. Somehow, a listener is able to overcome the phonetically disparate productions of phonemes to arrive at what a talker intends to say. (2006: 2382)

In this sense, some variation is also to be expected in pop songs, as the acoustically identical production of the same segment is nearly impossible (cf. Pardo 2006, 2382). This can be transposed from language to music because when a song is changed to an extent the audience cannot recognise it anymore, the point of the performance somehow gets lost. If there were too much variation, or variation that hinders perception, a song would lose its meaning and leave the audience bewildered. Just consider an instance where a singer forgets the lyrics of a song and substitutes the words for *lalala*. This surely happens many times, yet an audience expects to hear the words they know from the record, and paying someone to hear them sing *lalala* for three minutes is surely not what the majority of concert goers has in mind.

A live performance also differs from a recorded song version because it is not limited to the acoustic level alone. A live concert adds several layers to the performance of a song, such as the crowd watching the singer, gestures and facial movements, lighting, postures etc.; the audience singing along, clapping and cheering all play into the perception process, too. A song when performed live functions differently than one which is recorded in the studio.

It cannot be repeated, the level of attention within the listeners is generally lower because they are distracted and the purpose is simply another one when listening to a record. When going to a concert, one goes to "see" a band, not to hear them, which can be done quite satisfactorily at home. The expectations of the audience are based on what they heard on the record, but additionally hope that a live performance exceeds these expectations. A band is aware of this, and thus their aim to please the audience may not involve producing the best acoustic result.

7. Discussion

When talking about phono-stylistic variation in popular music, one should not forget to see the whole commercial product a pop song clearly is. If the pronunciation in pop songs were treated like the one in spoken language, one would leave out the whole idea of a song being performed. It is because of a song being performed in different settings, live and in the studio, that phono-stylistic variation is to be expected.

Stress placement and prosody is a difficult aspect of songs. The speech melody is outweighed by the melody of the song and thus investigating intonation as in spoken discourse would redundant. Also, stress has to be placed according to the rhythm of the song and can lead to syllables being stressed which usually would not be the case. Fortition processes may be expected to be distributed according to rhythm and also rhyme. In order to make two lines rhyme, words can be changed slightly in order to make a rhyme. An example for this occurs in the first verse of "Jacqueline", where *memory* is sung as ['mɛ̃ːmɒɹiː] in order to amount to the same number of syllables and rhyme with *shame to see*. In order to do so, two vowels are lengthened and one vowel is not reduced. This kind of processes do not

normally occur in spoken language but are typical for songs, where alternations are needed to fulfil the pattern of the songs.

Before analysing the data, some predictions about which processes occur can be made according to the theories outlined above. The notion of "ease of perception" has to be considered strongly, maybe even more strongly than with spoken speech, because when singing one requires more air and to distribute that successfully on all phonemes is at times difficult. It is hence to be expected that in songs with a high tempo and a high word rate phonemes will be strongly reduced or elided. It should also not be forgotten that emotions are conveyed with the singing and may have an influence on the pronunciation. It can be predicted that lines in songs which convey an aggressive meaning, will show traces of forceful articulation, such as strong aspiration.

Another factor is the "ease of perceptibility". In order to make an audience understand the words one sings, clear pronunciation is required. This should especially include words in prominent position or those which carry important information. As the chorus often is the most prominent feature of a song, it is to expected that the words in it should be marked by foregrounding processes more than maybe those in the verses.

It has to be noted, however, that the quality of perception is the singer's choice. An audience, apart from listening attentively, cannot change the speech signal and is totally dependent on the singer. If a singer, for whichever reason, chooses not wanting to be understood, perception is rather impossible. There are singers who sing just clearly enough to be understood on the records but when performing live the listener faces an unspecified sequence of sounds which simply make no sense, or only make sense to the experienced listener who already knows the lyrics.

Another very important aspect is whether the words that are sung in the live version are actually intended to be the same as in the studio version. There are some instances where the lyrics are modified in a live version, either by insertion or deletion of words, or by changing whole lines and verses. There are no rules for this. One has to be careful, though, not to compare segments and pointing out their differences when the intention was never actually to same.

For obvious reasons the scope of this paper is such as not to allow the investigation of all these details in full. Rather, it focuses on the most significant and salient aspects of the phenomena mentioned. There are many questions raised here which cannot be answered due there not being enough, if any, evidence in the collected data. This, however, leaves room for further investigations in the future, which cannot and will not be the issue here. Meanwhile, the collected data does allow some conclusions to be made about phono-stylistic variation in pop songs when performed in different settings.

7.1. Foregrounding processes

7.1.1. Foregrounding processes in consonants

7.1.1.1. Devoicing

Devoicing in this analysis only considers phonemes that would normally be voiced, this means devoicing after tautosyllabic segments as in *that's* has not been considered as this is an obligatory process in English (cf. Donegan & Stampe 1979: 142) and there have been no violations regarding this in the songs. Two examples of devoicing in the songs are the final /d/ in *kind*,

realised as [kʰãɪñd] (in "All I Wanna Do (Is Be With You)") or also in intervocalic position in *lazy*, which is sung as ['leɪz̥eə] in "Ice Age".

The voicing of consonants requires more effort from the speaker[12] than the production of a voiceless counterpart and consonants in any position (more frequently in initial and final position) are scarcely fully voiced, i.e. for the whole duration of the phoneme (cf. Roach 2000, 34f). This may be why devoicing proved to be a very frequent process in all the recordings. Even though from the perspective of ease of articulation, devoicing should be labelled a lenition process, yet in the context of pop songs, a devoiced consonant is easier to perceive than a voiced one, especially as vocals on music recordings have to stand out from a loud background which is filled by instruments. It is therefore that devoicing was here categorised as a foregrounding process.

There is a total of 83 devoiced consonants in the recorded versions compared to 74 instances of devoicing in the live performances, adding up to a difference of nine. The numbers of devoicing occurring in each individual song are given in table (3) below, and it has to be stated that only consonants which were completely devoiced, which at no time were voiced, have been taken into account.

There is a striking inconsistency in terms of whether devoicing occurs less often in a song when it is performed live than when performed in the studio. Apparently devoicing occurs overall less often live, yet in two songs, "All I Wanna Do (Is Be With You)" and "Sexy In Latin" it occurs more often in the live version.

[12] For a detailed discussion on the articulatory effort being higher in voiced consonants cf., for instance, Ohala 2005: 2.

(3)

DEVOICING	STUDIO	LIVE	DISCREPANCY
All I Wanna Do (Is...)	19	23	+4
Ice Age	29	23	-6
Jacqueline	14	4	-10
Sexy In Latin	7	11	+4
Tell Me What It's ...	14	13	-1
TOTAL	**83**	**74**	**-9**

The greatest discrepancy can be found in "Jacqueline" where there are 14 devoiced segments in the studio version compared to 4 in the live version. When looking closer at the transcription of the song, a striking pattern can be found as to which segments are devoiced. As devoicing does not only reduce the difficulty of the speaker to produce a sound, but also at the same time enhances perceptibility, this requires that in that specific location the voicing is not phonemic and the whole word or segment may not be confused with another, unvoiced one.

There are 11 cases of devoicing in the studio version of "Jacqueline" which involve either /d͡ʒ/ or final /z/ and it is exactly these eleven instances where the processes is not applied in the live version (there is an additional devoicing in the live version compared to the studio version, which adds up to the total of -10). In the studio version final /z/ is sung as [z̥] five times where [z] is sung in the live version, while /d͡ʒ/ becomes [d̥ʃ] six times where [d͡ʒ] is voiced in the live recording. It is essential to notice that most devoicing processes which involve either the final voiced fricative /z/ or the voiced affricate /d͡ʒ/ do not need to be voiced for the phoneme intention to be understood. For example, devoiced [d̥ʃ] only occurs in the word and title *Jacqueline.* Perception is not influenced negatively (as in rendering it intelligible) because there is no word in the English language, which is

comprised of the sound sequence /ˈʧʌkəliːn/ and confusion is not possible, and the listener expects /ˈʤʌkəliːn/ because of the song title anyhow. Hence compromising voicing and thus making production and perception easier seems logical.

If production and perception are simplified by devoicing, why then are there more devoiced segments in the studio version where breathing is easier, recording can be done in smaller sequences and the surrounding noise is lower? Finding that devoicing is only more frequent in two of the live versions was rather unexpected. This may be related to singers trying to sing more clearly in the studio, after all, this is the recording listeners will hear over and over again, and if the words can only be understood by the tenth time one hears it, it makes the lyrics seem rather redundant. Also, voiced segments "sound more" as in they are more suitable for singing because only voiced phonemes can carry a melody. Additionally, one has to remember that devoicing, especially of final voiced elements, while being a natural process in English, is suppressed at an early stage. The fact that the process is applied, meaning final voiced obstruents are devoiced, is indeed quite startling.

There are, however, two songs where more segments are devoiced in the respective live versions than in the studio recordings. In one of them, Paul Weller's "All I Wanna Do (Is Be With You)", word inital /d/ and /b/ are devoiced in the very frequent words *be* and *do* because the are used in the chorus. Especially towards the end of the song, devoicing is very prominent in the live version and in two very emphatic instances during the climatic ending /b/ becomes an ejective, [b̥'] . The foregrounding effect the devoicing has is clearly related to their prominent position in the song. Seeing as the words of the chorus are exactly the same as those of the title, fortified articulation may be expected and indeed happens. In a live setting a sharper

contrast to the surrounding segments is vital to the understanding of the words. By devoicing consonants, this foregrounding effect can be achieved.

7.1.1.2. Aspiration

Aspiration is a "[...] burst of noise [...] a period during which air escapes through the vocal folds, making a sound like **h**." (Roach 2000: 34). As aspiration is default in fortis consonants in initial position of stressed syllables it is by no means unexpected that it is the most frequent process which was applied in the songs; the total number of occurrences in studio recordings is 116 compared to 92 live, which is 24 less. The most frequent position of aspiration is, as mentioned above, word initially and mostly affects the stops /t/ and /k/, yet these are also aspirated in medial position in unstressed syllables and in word final position: in "Sexy In Latin" *back* is realised as [bak^{h}] with the final consonant being aspirated. As table (4) shows, the occurrences of aspiration are generally higher in the studio recordings than in the live recordings.

(4)

ASPIRATION	STUDIO	LIVE	DISCREPANCY
All I Wanna Do (Is...)	21	18	-3
Ice Age	36	20	-16
Jacqueline	10	5	-5
Sexy In Latin	20	20	0
Tell Me What It's ...	29	29	0
TOTAL	**116**	**92**	**-24**

Two songs have exactly the same number of aspirated segments in the live and studio versions, "Tell Me What It's Worth" and "Sexy In Latin", which,

however, does not mean that exactly the same phonemes were aspirated, but merely that the total number coincidentally turns out to be the same. The biggest discrepancy can be found in "Ice Age" where in the studio aspiration was applied 36 times and only 20 times in the live performance. All instances of aspiration in the studio recording occur in a default position, where it may be expected, but 16 less of those are realised in the live version. There are two factors which may explain this difference:

First of all, aspiration requires quite some extra air, which can be a problem when singing live and not having the best breathing technique. Quite contrastingly, in the studio, one can record each part of the song, each line, separately and there is sufficient time for breathing. What enhances the difficulty of aspiration in "Ice Age" is that the song is sung faster live than on the record, which makes the pauses for breathing even shorter and articulation is difficult enough without aspiration. In the studio recording, aspiration also underlines the staccato of the song. Most words are monosyllabic and repeated several times in total accordance with the beat, and each word stands on its own and is clearly separated from its neighbours. To summarise, a higher tempo allows for less breathing time and hence aspiration occurs less often

And secondly, as pointed out in section 4 on Emotional Phonetics above, aspiration can also express emotional states and is not only used as a foregrounding process. Both in the live and the studio recording of "All I Wanna Do (Is Be With You)" there is a passage which differs greatly from the rest. To draw greater attention to this part of the song, it is marked by different instrumentation, as especially the guitars are more rhythmical while in the rest of the song they are more melodic which gives this passage a completely different, more aggressive mood. The aspiration in the words "I bathe in dust / I feel the most / I twist and turn turn / I'm lost and found" is

comparatively heavy, especially in final position in *dust* and *most*. The only instance where the live version differs in points of aspiration is in *most* which is reduced to [mɒs], compared to [məʊstʰ]. The reason for this passage being quite similar in the studio and in the live recording can be explained resorting to Emotional Phonetics, as the data suggests that agitation leads to higher intensity of the speech signal, and this includes aspiration (cf. Scherer 1986: 161).

Additionally, as already touched upon in section 4, in the word sequence "kill, kill, kill" in "Tell Me What It's Worth" the initial consonants are aspirated heavily. There the aspiration can likewise be attributed to emotional expression. The aggravated feelings, which were even audible on both recordings and thus not only perceptible to listeners present in the studio or the venue, are not only reflected in the words themselves but also in their forceful articulation as [kʰɪɬ kʰɪɬ kʰɪɬ][13].

There are only few instances where aspiration occurs in word final position, presumably to make the final consonant easier to perceive. In "Tell Me What It's Worth", final /k/ is aspirated in three words in the studio version and two of those also in the live version. The same phoneme is aspirated twice in *park* and *dark* in "Sexy In Latin", yet only in the live version. Both words form a rhyme of two lines at the beginning of the second verse and the singer apparently placed comparatively more stress on the words when performing live than when in the studio. When looking at table (4) one can see that the numbers of aspiration for live and studio recording are the same, which then means that in two instances aspiration in default setting were dropped in the live version, where it was in the studio version.

[13] It is not always possible to interpret the singer's emotions quite as clearly as in this passage, as judgement on emotions is difficult on vocal cues alone. In this passage, the succeeding line "when everything starts to suck" in addition to the vocal cues allowed for an interpretation that is not necessarily possible elsewhere.

In “All I Wanna Do (Is Be With You)” two final stops are aspirated, in *mind* and *kind*. It is interesting to notice that the stops involved were probably intended as voiced consonants, were devoiced and then aspirated. Aspiration of a devoiced stop usually makes the listener perceive it as its voiceless equivalent (cf. Roach 2000, 34) yet in word final position, where aspiration does not generally occur in English - with exception for careful speech where speakers want to foreground certain elements - this aspiration does not lead to the words being interpreted as **maint* and **kaint*, which do not exist in English.

7.1.1.3. Affrication

The process of affrication involves a stop to which a fricative is added to form an affricate. Affricates thus are “rather complex consonants. They begin as plosives and end as fricatives” (Roach 2000: 48). This then means that if a stop is affricated this makes the sound more complex and hence more difficult to produce. Based on this it is here classified as a fortition process because affrication makes a segment stand out from its environment.

Yet not all affrications in the songs can be seen as fortition processes, but may rather be classified as lenition processes. Why is this discussed in the foregrounding section then? - Because what actually happens is more complex. This type of affrication involves the consonant cluster /dɹ/ which is realised as [d͡ʒ] or [d̟͡ʃ] in the words *drinking* (in “Sexy In Latin”) and *drowning* (in “Tell Me What It’s Worth”). The result is clearly an affricate and to arrive at this output more than one process has to be applied: The environment to form an affricate is suitable as both consonants /d/ and /ɹ/ are homorganic, a precondition to form affricates (cf. Roach 2000: 49). The voiced alveolar approximant [ɹ] is then changed into a postalveolar fricative which actually eases pronunciation in this context. The output so far is then

[d͡ʒ] which, however, is only realised once - in the studio recording of "Sexy In Latin" - out of the four times this type of affrication happens. The other three affricates succumb to devoicing as well and, this being a foregrounding process (as established above), counteracting the backgrounding process of the actual affrication.

It is questionable as to whether the realisation of /dɹ/as [d͡ʒ] or [d̥͡ʃ] is easier to understand by an audience or not. The affrication only occurs in two songs and there only once, and the number of possible environments for this type of affrication is very small. In fact, there are in all five songs only four instances of /d/ which is followed by /r/ and the affrication process is applied in half of those. Out of two possible applications in "Tell Me What It's Worth" one is realised, the one possible occurrence in "Sexy In Latin" is realised, and in "All I Wanna Do (Is Be With You)" affrication does not happen. It is apparent that the process is dependent on the singer and even then is not applied consistently.

A different affrication process which (in the following case) is somehow similar to aspiration as it here involves more air to be produced can be heard in "Jacqueline". In this song, the medial stop of the name *Jacqueline* is sometimes not realised as [kʰ] but as [k͡x], which means that instead of aspiration a fricative follows the stop, resulting in an affricate. There are three instances of this in the studio version compared to one instance of [k͡x] in the live version, yet there it occurs in a different word in the first verse, *kiss.*

When there is affrication of this type in a song it may be expected that many of the voiceless velar stops, which are not affricated, are aspirated. This holds true for the studio version of "Jacqueline" where there are 10 aspirated segments, but in the live version only half as many; 3 of those plus the one occurrence of affrication are gathered in the first verse. The position in the

first verse may be attributed to the fact that a singer, when not sufficiently trained, may have trouble breathing as the song progresses, as I already pointed out in my hypothesis. Hence, foregrounding processes are expected to occur less often in live recordings.

While one plausible explanation for this type of affrication only occurring in only one song, and especially this one, is that the band performing the song, The Coral, are from Liverpool, where the affrication of voiceless stops forms part of the dialect (cf. Hughes, Trudgill & Watt 2005: 98). It is interesting, however, that this foregrounding process almost vanishes in the live version and a less regionally marked articulatory target is chosen. Again this may be attributed to respiration restrictions, which, however, contradicts another distinct process in this song. A more plausible explanation, therefore, may be that the performance in front of a London audience led to the singer choosing a different realisation[14].

There are also occurrences of the voiceless labial-velar fricative /ʍ/, which is a phoneme whose realisation is again limited to this song. The phoneme occurs twice in the live version and thrice in the studio version at the beginning of the interrogative pronouns *what*, *where* and *when*, yet not at the same position in the song. As this kind of phoneme in this song only occurs in inital position of *wh*-words (and in all other songs not at all), it is necessary to point out that throughout the song there are only 6 possible positions for /ʍ/ to be voiced, and the percentage of its realisation is quite high (50% studio and 33% live). According to Hughes, Trudgill & Watts (2005: 97-99) this phoneme is not part of the Merseyside dialect but is generally attributed to more northern accents. It is, nonetheless, clearly a foregrounding process and even if its use can not be attributed to a dialectal input, it makes the

[14] Trudgill's study on pop songs included accommodation theory as an explanation for phono-stylistic variation in respect of an audience. While it may not account for all variation, here, it appears to be suitable.

segment stand out and easier to perceive. In terms of respiration as mentioned above, similar to aspiration, the spirantisation and simultaneous devoicing of the labial-velar approximant /w/ can be predicted to occur less in live situations but the set of data on hand is too limited to provide conclusive evidence.

7.1.2. Foregrounding processes in vowels

7.1.2.1. Lowering

The process of lowering involves vowels to be produced with the mouth being open wider and has the effect of more sonority and thus loudness (cf. Donegan 1978: 35f.). In terms of making one understood, it seems reasonable to assume that processes which make vowels more easily perceptible are applied in singing[15]. The stronger the deviation from the neutral/central position, the easier to identify a vowel becomes, and in a live recording this may be essential. In "Ice Age", for example, the vowel in *far* is lowered in the live version to [faː], compared to the studio recording [fɑː]. As argued, the vowel becomes more sonorant because of the lowering process and thus is easier perceptible.

There are several instances of lowering in the songs, which here means that one realisation is lower than the respective one in the other recording. As predicted, lowering occurs more often in the live versions than in the studio versions. One instance is the case of *crunching* in "Tell Me What It's Worth" (discussed further in 7.2.2.) where in the studio the first vowel is realised as [ɤ] and live as [ʌ]. In the live version the lower vowel was chosen and fulfils listeners' expectations more than the closed variant. The realisation as [ɤ] is rather surprising as the singer, Devonte Hynes, is from Southern England and his singing throughout the song is generally Southern,

[15] This does exclude, of course, singers who mumble intentionally to obscure their lyrics.

as he clearly distinguishes between /ʌ/ and /ʊ/. When performing live, however, a more listener friendly (and more consistent) realisation is chosen. Listener friendly in the sense that a London setting, with a presumably Southern audience, will be more used to a Southern realisation.

In "Sexy In Latin" there are four different segments which are affected by lowering in the live version. First of all, the diphthong /əʊ/ in *hold* and *don't* is realised as [ɐʊ], with the first element noticeably lowered. This does not affect all instances but the lowering affects diphthongs at the beginning of the song. This instance of lowering, which affects the first element of the diphthong /əʊ/, enhances sonority and thus perceptibility especially in a song with quick pace and loud instrumentation

Another apparent instance of lowering in "Sexy In Latin", which occurs in the word *scratched* in the first verse of the live version, cannot be conclusively identified as a foregrounding process. In the studio recording the word is realised with an [æ], and in the live recording the vowel is lowered to [a]. In the third verse of both versions the word is sung as [skɹat͡ʃ] with the lower variant. The lowering of [æ] to [a] cannot only be explained as a foregrounding process. It could be judged to be just that when looking at the data alone and not considering that the singer is from the North of England where the realisation of [a] is common in words like *scratch* (cf. Wells 1982: 353-356). The difference here is hence not caused by a foregrounding process but merely by the choosing of a different articulatory target. In the one instance of [æ], apparently a different, more Southern target was chosen. It seems strange that this happens in the studio version and not in the live version when facing a Southern audience. Yet again the [a] might be consciously used to mark the singer's Northern identity and to contrast himself from the audience.

The similar change of input can be heard in the word *come* where in the studio versions the vowel is [ʊ] and in the live version is lowered to [ʌ]. Interestingly enough here the inputs swap place and the Southern articulatory target is chosen for the live version. This may be due to the fact that the Southern variant is more sonorant or that the performance was in London, hence in front of a Southern audience. Another example in which the Northern target outdoes the Southern target used in the recorded version is *dark* where the vowel in the live recording is [aː] compared to [ɑː]; yet again the fronting leads to more sonority. Generally it can be said that even though the singer is not dialectally consistent all instances of lowering enhance sonority and with that intelligibility.

Also, seen in a larger context, the more back vowel [ɑː] seems to be out of place. *Dark* forms a rhyme with *far*, and as *far* is realised with the front open vowel [aː], the more Southern realisation of *dark* does not rhyme fully. In the live version the rhyme works perfectly.

7.1.2.2. Diphthongisation

Diphthongisation is a process by which pure vowels are lengthened and a second element is added to form a vowel glide, a diphthong. In England the long vowels /uː/and /iː/ tend to be diphthongised in the South, this happens in Received Pronunciation as well as in Estuary English (cf. Gimson 1989: 102), making this a rather common process. Compared to the other process which were analysed, diphthongisation turns out to be applied rarely. As table (5) shows, the process is generally applied more in studio recordings than in their respective live version.

(5)

DIPHTHONGISATION	STUDIO	LIVE	DISCREPANCY
All I Wanna Do (Is...)	1	0	-1
Ice Age	7	3	-4
Jacqueline	0	0	0
Sexy In Latin	2	2	0
Tell Me What It's ...	2	0	-2
TOTAL	**12**	**5**	**-7**

There is a startling amount of diphthongisations in "Ice Age", where the occurrences add up to 7 in the studio and 5 live, making up more than half of the total. In the studio recording of the song, the word *reading* is consistently realised as [ˈɹɪə̯ɗɪn] with the long front close vowel /iː/being diphthongised. The live version differs from this, and the diphthongised variant is only sung once.

The remaining instances of diphthongisation occur at word final position in *sorry* and *lazy*. In both cases the final vowel is lengthened and diphthongised to [eə], and again the process is applied more often in the studio version. In the live version, which is played quicker than the studio version, diphthongisation may not be possible because the time span to perform the vowel glide is too short. The monophthong in *reading* on the live recording is audibly shorter than the diphthong on the studio version and it may be due to the tempo that diphthongisation is not applied.

7.1.2.3. Lengthening

When an otherwise short vowel is lengthened, this is usually connected to it taking a more prominent position in a phrase, for instance when carrying

stress (cf. Donegan 1978: 54). A prominent position in a song is towards the end of a line, which is often part of a rhyme. It comes as no surprise that lengthened vowels are found at the end of lines in the songs observed here.

To form a rhyme with *see* the final vowel in *memory* in "Jacqueline" is lengthened to [iː] in both versions. And in "Sexy In Latin" realisations of *you* vary between [jəˑ], [jɜˑ] and [jɤː]. When *you* is sung at the end of a line it is first of all reduced to have a more central quality but then lengthened. It is, however, never realised as [juː]. This makes it, in a way, a foregrounding process applied to a backgrounded segment. The longer central variant may still be easier to produce than the one with the close back vowel.

7.1.3. Insertion

The lengthening of syllables or phonemes is often achieved by the insertion of an extra element. If syllables have to be stretched out over more beats or notes then insertion is usually applied. This can mean the insertion of an extra vowel, as in "Sexy In Latin" where *on* is realised as ['ɒˈhõn] to fit to the rhythmical pattern of the song. Another, and altogether different example can be found in "Ice Age" where at several occasions the word *and* is inserted to fill the gap between two lines. The numbers of insertion are different from song to song as table (6) shows and are largely dependent on the song's individual needs.

(6)

INSERTION	STUDIO	LIVE	DISCREPANCY
All I Wanna Do (Is...)	0	2	0
Ice Age	3	7	+4
Jacqueline	1	1	0
Sexy In Latin	14	15	+1

INSERTION	STUDIO	LIVE	DISCREPANCY
Tell Me What It's ...	1	4	+3
TOTAL	**19**	**27**	**+8**

When a word has to be stretched over a longer period of time, to cope, an additional, identical syllable is inserted. So, for instance, in "Ice Age" *have* is sung in a way to last for additional time and realised as [ha'haːv], where the first two phonemes are sung twice and the vowel is lengthened additionally. This is similar to what happens in "Sexy In Latin" where towards the end of the song the words *hold on* are repeated six times, and each time sung as [həʊhəʊɬd ɒ.ɒ̃n] with a syllable inserted in each word. The purpose of the insertions is simply to make the words carry more melody and these instances alone explain the high numbers of insertion of the song compared to the other ones.

The song where the discrepancy between studio and live versions is the highest is "Ice Age". There are four more insertions in the live versions compared to the studio versions, and all four are of the same type. As mentioned above, the word *and* is inserted at the end / beginning of lines to link them and to fill the gap. As the lyrics are written to have one syllable for each beat the *and* fills up empty space and makes the singing more continuous in the live version. As the insertion of the word does not have any value in terms of meaning but is only restricted to its function of filling a rhythmical gap, this can be attributed to phonology, and is similar to attaching an extra vowel, as in "Sexy In Latin".

In the same song a different instance of insertion can be found. In the first line, the word *one* is sung in the studio recording as ['wɒ̃nə̆h] and in the live recording as ['wɒ̃nəɦ]. In both cases a schwa and a glottal fricative have been added to the end of the word. Again the syllable is added to make a

monosyllabic word to fill the space where two syllables would fit according to the rhythm of the song. The realisation of final [h] and even final [ɦ] is, however, very uncommon in English. Especially the latter phoneme is rather unexpected as it generally does not exist in English, while the additional voiceless counterpart, [h], seems to appear in preparation of the preceding /h/ which follows in *who*. The use of the voiced phoneme is probably related to the fact that voiced consonants can carry melody, and as pointed out in section 7.1.1.1., devoicing is less favoured in live performance.

Another aspect of insertion is the use of *intrusive* /r/ in "All I Wanna Do (Is Be With You)". There, the only instance of such an /r/ occurring in the whole set of data (it is admittedly the only environment where this could happen) is between the words *draw upon*. In both the live and the studio version those words are sung as [dɹɔːɹ‿əˈpɔ̃n] and are clearly linked by the use of an /r/. In effect this makes the passage more fluent and sound "more gentle" as there is no break. A sharp contrast to this can be found in "Ice Age" where an *linking* /r/ between the words *for a* is not realised and each word stands separately. It has been pointed out above that in this song each syllable occupies one beat and by not linking the elements this effect can actually be achieved. Additionally, both songs differ in their moods and a more "gentle" pronunciation in "All I Wanna Do (Is Be With You)" includes intrusive /r/ while the more aggressive singing in "Ice Age" excludes linking by an /r/.

Differently to what Trudgill found in his 1982 paper, the singing of the bands analysed here is consistently non-rhotic. There is only one non-prevocalic /r/ to be found in the data, in "All I Wanna Do (Is Be With You)". In the live version *turn* is twice realised as [tʰɝ̃ːn], while in the studio version this remains completely non-rhotic.

7.2. Backgrounding processes

7.2.1. Voicing

The process of voicing affects elements that are usually realised without voicing but in a specific environment are voiced. In the songs this generally involves the phoneme /t/ in intervocalic position. As mentioned in section 2., intervocalic /t/ voicing was an element inherent to "pop-song style" and apparently deemed to be very common - according to Peter Trudgill (see section 2). Even though there are some instances of /t/ being voiced, which is also sometimes called "flapped", it cannot be said to be the norm. A voiced /t/ was represented in the transcription as [t̬] rather than [d] or [ɾ] in to be more iconic.

As it is shown in table (7) the process of voicing is applied rarely in some songs, and the overall numbers are fairly small, 28 studio and 26 live. The number of intervocalic /t/which could be voiced is much higher.

(7)

VOICING	STUDIO	LIVE	DISCREPANCY
All I Wanna Do (Is...)	5	5	0
Ice Age	1	1	0
Jacqueline	3	3	-2
Sexy In Latin	6	8	+2
Tell Me What It's ...	13	11	-2
TOTAL	**28**	**26**	**-2**

In those songs with more than one instance of /t/ voicing, the process usually occurs in the same sequences. "Tell Me What It's Worth", for instance, has such a high number of voicing processes because in each chorus

the intervocalic /t/ in *what it* is realised as [t̚]. This realisation is constant throughout the song. The discrepancy of two less applications in the live version are due to the fact that one sequence where the process is applied in the studio version is simply not sung in the live version, that is, the words are different. The other instance of voicing which is not realised in the live recording is the voicing of the final /k/ in the very first word of the song, *crack*.

7.2.2. Monophthongisation and shortening

In the recordings, some diphthongs are reduced to monophthongs. The process of monophthongisation describes the realisation of an otherwise diphthong like, for instance, /aʊ/, as a monophthong like [ɑː]. Overall the number of monophthongisations is larger in the studio recordings than in the live recordings, but, as table (9) shows, there are actually three songs where the process is applied more frequently in the live version.

(9)

MONOPHTHONGISATION	STUDIO	LIVE	DISCREPANCY
All I Wanna Do (Is...)	10	11	+1
Ice Age	23	28	+5
Jacqueline	13	7	-6
Sexy In Latin	12	8	-4
Tell Me What It's ..	4	5	+1
TOTAL	**66**	**59**	**-7**

There are not many instances in the songs where monophthongisation occurs alone without a simultaneous application of shortening. These instances can be found in the songs of the Northern bands, Little Man Tate and The Coral. In their singing some diphthongs are realised as monophthongs because of different dialectal articulatory targets and are not shortened.

In the live versions the process is applied less than in the studio recordings, so, for instance, *came* in "Jacqueline" is in the studio [kʰɛ̃ːm] and live [kɛ̃ĩm]. There are more instances like this, and it seems as if the singer opted for the non regionally marked variant when performing live more often than in the studio.

A different type of monophthongisation occurs frequently in "Ice Age" and "All I Wanna Do (Is Be With You)", where *I* is reduced to the first part of the diphthong /aɪ/ to only [a]. The numbers for "Ice Age" appear to be rather high but this is simply because *I* occurs so many times in the song and is in the majority of cases reduced.

Generally it can be said that when monophthongisation can be explained through a different dialectal input the occurrences are more frequent in the studio versions. If, however, segments are monophthongised and reduced to ease articulation they occur more often in the live versions.

7.2.3. Elision

The most frequent backgrounding process in the songs was elision, a process the nature of which Peter Roach sums up to "under certain circumstances sounds disappear." (Roach 2000: 142). This chapter will point out some of these "circumstances" of elision in the songs and try to explain why and which phonemes were made to "disappear".

But first returning to the numbers, there were a total of 95 phonemes elided in the studio recordings and 110 in the live versions, summing up to 15 more in the latter. Table (8) illustrates the distribution of elision processes for each song performed live and in the studio.

(8)

ELISION	STUDIO	LIVE	DISCREPANCY
All I Wanna Do (Is...)	10	15	+5
Ice Age	15	23	+8
Jacqueline	10	10	0
Sexy In Latin	35	32	-3
Tell Me What It's Worth	25	30	+5
TOTAL	**95**	**110**	**+15**

Overall, elision occurs mostly in final position and it is mostly alveolar consonants that are elided. So, for instance, the final /t/ of *must, just* or the final /d/ in *and* are not realised at all. Another common elision is the loss of /v/ in final position, as in *I've*, which is either realised as [aɪ] or in a even more reduced form: [a]. The elision of final elements happens also in prominent position, in "Tell Me What It's Worth" *worth* is an essential component of the chorus, yet the final fricative /θ/ is not once realised, rather, articulation stops after the long central vowel.

Yet, in this song, elision not only occurs word finally, but also in inital consonant clusters, as in *crunching*. The word thus becomes [ˈkʰɤ̃nt͡ʃɪ̃n] in the studio recording and [ˈkʰʌ̃nt͡ʃɪ̃n] in the live recording with the voiced alveolar approximant being dropped in both versions. Similarly *clean* is sung as [kʰĩˑn] in the studio version, while the /l/ is realised in the live version. Contrary to elision of final /d/ or /t/, these cases obscure understanding.

Elision within a consonant cluster of that sort is not a common backgrounding process, at least I have not come across this before, and to guess the speakers intention from this output is rather difficult, if not impossible[16]. When listening to the recorded version and simultaneously reading the lyrics one might not even notice the elision, but when listening to the song unaided the difference is quite startling. This may be a reason why the elision does not occur in the live version.

Taking a closer look at the one song where less elision occurs in the live recording than in the studio recording, "Sexy In Latin", some interesting aspects present themselves: in the first line of the song the two words *village store* are realised very differently live than on the record. Especially the affricate sequence that links the two words is affected by lenition processes. While in the studio version the passage is sung as ['vɪlɪd̥ʒ ʃtɔːɪ] with the initial /s/ being assimilated to the preceding 'voiceless' affricate (see section 7.1.1.1. for the devoicing process regarding the affricate), the consonant cluster is still realised. In the live version the relevant words are realised ['vɪɬɪʃ sɔːɪ] with two alveolar stops elided. While in the studio version one backgrounding process, assimilation, occurs, in the live version a completely different one changes the sequence: elision.

In terms of perception, the application of both processes at the same time would be impossible. A supposed form of *[vɪɬɪʃ ʃɔːɪ] which combines assimilation and elision may not be understood. The question now is why the singer chose two very different lenition processes for the same sequence in two different situations. The most likely explanation lies with the tempo of the song. The studio version, even though being upbeat and demanding

[16] I asked three people to listen to the relevant passages in the studio recording and tell me what they understood. While two of them offered no guess for *crunching* but could tell me what they heard was ['kʰɤ͡ʧɲ̩] one guess was *kitchen* as it was deemed to be the word closest to this output. The other item, *clean*, was understood by all three as *can*, and the sounds they heard were describes as [kʰə̃n].

rapid articulation of the two words, is still considerably slower than the live performance. While this seems to have generally little effect on the pronunciation in this song, the stops in this cluster appear to have been elided for the sake of speed, as it would be difficult to articulate the cluster in time to fit the rhythm of the song. This type of backgrounding process is fairly common in rapid speed generally and to elide elements in this difficult cluster eases production noticeably.

Another interesting elision occurs in "Ice Age". Here, in the the first verse *news*, in the studio recording, is sung as [nuːz̟] which in the live version this becomes [njuːz̟]. The question is whether there is an extra voiced palatal approximant inserted in the live version or this phoneme is elided in the studio version. Judging from the other examples of the song where *nudity* is constantly realised as ['nju̟ːdɪtʰi] in both versions it follows to classify the process as an elision which means that the segment is backgrounded in the studio version. By doing so the perception for the listener is less simple yet, and this is another important aspect, the elision leads to *news* rhyming better with *truth* in the preceding line. The quality of the rhyme is given up in the live version for better understanding.

Another form of elision, which is very common in all forms of casual speech, is the reduction and cliticisation of very frequent words, and especially the verbs *have* and *be* are affected. *Have* is either reduced to [əv] or merely [v], where the former appears frequently in "Sexy In Latin" and the latter in "Ice Age". Only in stressed position *have* is fully pronounced, namely in "Jacqueline" where *has* is sung exactly on the beat is hence the most prominent syllable in the line. Therefore reduction is neither applied in the studio nor in the live recording. In the case of *be*, the first person singular *am* is in all instances reduced to [m] and in many instances the third person

singular is realised as [s]. Because the verbs do not carry stress or actual meaning, they are strongly reduced.

Some syllables are reduced even further up to a point where a consonant is then syllabified and stands for the peak of the syllable. The process can be found in "Sexy In Latin" where in the both versions *changing now* is reduced to [t͡ʃẽĩnʃn̩ n̩] with two neighbouring nasals being syllabified. The syllabification here allows for all the words to fit the line according to rhythm. As mentioned before, the tempo of the song is relevant in this respect, and as the liver version is quicker than the studio version, syllabification occurs more frequently. In the live recording *written* is constantly realised as ['ɹɪʔn̩] while in the studio recording this is ['ɹɪʔĩn]. The words of the chorus are " Something's happening / It's written on the wall that you're sexy in Latin / This time come on / Everything's changing now we can't hold on" and each section between slashes is sung to fit one bar. The first line consists of four syllables and the next line, which is realised in exactly the same amount of beats, thirteen syllables are sung; similarly in the third line four syllables are followed by ten in the fourth line. For the singer to be able to articulate all of these syllables in time, syllabification and other reduction and elision processes need to be applied. More syllabification occurs in the live version because it performed more rapidly.

7.3. Location, location location

In the first verse of The Rifles's "I Could Never Lie"[17], singer Joel Stoker realises *have* as [æv] and drops the initial /h/. All throughout the rest of song, however, he realises initial /h/. This seems inconsistent and random,

[17] This song was not part of the actual investigation but similar to examples quoted above it helps to illustrate processes in pop songs in general.

yet it can easily be explained. In Natural Phonology the setting and phonological environment of a segment is vital to its realisation, and the pronunciation of pop songs depends on the music that surrounds it. The formerly mentioned *have* occurs at the beginning of the song, because instrumentation is limited (one guitar, tambourine and bass) and the background, and hence distracting, sound is minimal. A realisation without an initial /h/ does not affect perception at all, yet the word *hide* in the chorus, being accompanied by the full band, needs the initial fricative to stand out from the musical background and to be recognised. Also a realisation without /h/ is not unexpected as the singer is from London and the dialectal input surely does not include the initial fricative.

This example illustrates how important it is not to forget the position of a specific segment within a song. There are so many features which only occur once or twice in a song, but instead of labelling these as being inconsistent, one should consider the reason for why the segments occur in this realisation at this position. One example from the songs that have been under close examination in this paper shows the importance of positioning nicely. In "Sexy In Latin" by Little Man Tate, singer Jon Windle includes two very interesting fortitions in the last verse. It has to be said that this verse, contrasting to the pervious two verses, is accompanied more quietly and especially in the relevant part instruments are only strummed once a bar. This allows for the vocals to be heard very clearly and comparatively loudly.

In a live setting surroundings are different and while in the recorded version the passage shows no peculiarities, in the live version the word *something* is suddenly realised as ['sʊ̃mp͡fɪ̆n] with an (emphatic) bilabial stop inserted between the nasal and the voiceless fricative. This fortification process is similar to the one described by Donegan and Stampe, where a stop is inserted between a nasal and a spirant (Donegan & Stampe 1979: 143). It has

to be mentioned that this happens only once in ten times *something* is sung in the song[18]. Clearly the prominent position of the word affects its realisation with an emphatic stop inserted.

Another striking example, just a few beats earlier, should be considered as well. In the recorded version *twenty-four* is realised as [ˈtʰwẽni fɔː], a rather common variant, yet in the live version this becomes [ˈtʰwẽni f̰ʼɔː]. The first process which will be considered here is the elision of /t/ in the second syllable of *twenty,* a backgrounding process, which occurs in both versions. The more remarkable process happens in the word *four* in the live version, where the initial sound is a dentolabial ejective, which is a sound that is neither included in the English phoneme inventory nor in the IPA chart. This makes it an unusual finding, yet it is easily identified as a fortition process. By transforming the fricative into an ejective, the segment is foregrounded and contrasts more with its environment, especially as a backgrounding process is applied in the preceding segment.

Still it can be safely assumed that the singer did not intend to produce a dentolabial ejective as he is probably not aware of either this phoneme existing, or the fact that he is capable of producing it. The intended sound must have been merely the fricative [f] and indeed so must have been the perceived sound, yet the fortition/foregrounding process gave a different output. It is striking that two rather exceptional instances of foregrounding should occur within one verse, but it is less surprising when considering their prominent position in the song. The singer apparently tried to pronounce the third verse very clearly and distinctly, and by doing so he subconsciously fortified the voiceless dentolabial fricative.

---

[18] For the sake of completion, a similar process is applied in the second verse where, again in the live version, *far* becomes [p͡fɑː]. The prominent position of the word towards the end of the line, and it's function as a rhyming word may account for the fortition/foregrunding process.

When looking at the other songs, it came as a surprise to find two more ejectives. There is a alveolar ejective in "Tell Me What It's Worth" in initial position of *dead,* and in "All I Wanna Do (Is Be With You)" a bilabial ejective is realised in the word in *be* (also see section 7.1.1.1.). Similar to the dentolabial ejective in "Sexy In Latin" the positions of the phonemes are not coincidental. *Dead*, in the second verse of the song, is realised as [d̥'ed][19] in both the live and studio recording. The word, especially in this context, is emotionally loaded and choosing such a foregrounded realisation appears to be connected to this. The realisation is not dependant on the setting but on the content or meaning of the word which explains why the alveolar ejective is produced in both recordings.

In "All I Wanna Do (Is Be With You)" a bilabial ejective appears twice towards at the end of the live recording of the song in *be*. Here the musical context and setting influenced the fortition process and can be explained by the need of foregrounding the element because of the climatic effect at this stage of the song. The point of the song in terms of meaning is to "be with you" and at the end this is repeated six times in a very short time. The use of the ejective makes this wish, despite its obvious importance underlined by repetition, seem even more forceful.

7.4. Different lyrics

There are some differences in the live and studio recordings that go beyond phonetics and phonology. In some songs the lyrics are simply changed and different words are sung. A comparison between these passages is then

[19] The transcription shows the ejective as a devoiced [d̥], based on the nature of ejectives, they can never be voiced, however, in this context, the representation as devoiced instead of voiceless was chosen to be more iconic if not entirely phonetically correct.

redundant because they should differ. The fact that foregrounding and backgrounding processes are applied nonetheless cannot be doubted.

There are only two songs where the words of the live version differ from the one in the studio recording (leaving the odd insertion of *and* in "Ice Age" aside). Those are "Tell Me What It's Worth" and "All I Wanna Do (Is Be With You)", and in each song two lines are different. In "All I Wanna Do (Is Be With You)" *I'm not out to* is twice substituted in the live version by *I'm not looking to*. The lyrical change in "Tell Me What It's Worth" is the following: *So tell us that we're spelling everything wrong* becomes *tell us that we're doing everything wrong* and *'cos I've just turned bright red* is live *I think I turned red*. Notice that the live versions are each one syllable short.

8. Conclusion

The aim of this paper is to show the phono-stylistic variation in British pop songs, and the main question is whether the pronunciation differs in live recordings when compared to studio recordings. Then, given the existence of such variation, the subsequent question ensued how the pronunciation varies and how this can be explained with help of Natural Phonology.

The answer to whether there is variation in a singer's speech due to different recording settings can be answered with an irrefragable yes. Differences regarding pronunciation can be heard in the chosen recordings, sometimes even without observing it too closely. When considering the phonetic transcriptions of the songs, much variation is to be found and to be considered. One main goal was to attribute the variation to processes and then group them, according to the distinction as proposed by Natural Phonology, into foregrounding and backgrounding processes.

Five songs were transcribed and patterns of variation seemed to appear. Seven distinct processes have been chosen to be analysed in depth because they concerned all recordings. The application of these processes were then counted and the numbers were discussed in the respective sections. Even though the numbers were sometimes contradictory and the application of one process was more frequent in one song than others, some general remarks can be made.

The dichotomous concept between speaker and listener as proposed by NP can explain a great number of processes which determine the phonetic variation in the songs. There are some processes which favour the singer, as for instance, elision occurs more frequently in live recordings because it is demanded by the greater exertion on the singer's behalf. In a studio setting,

where a singer can re-record there is no such pressure and greater emphasis can be put on producing a result which can be understood by the listener. It therefore comes as no surprise that fortition processes like aspiration are applied more frequently in studio recordings.

It was shown, however, that the theory of Natural Phonology cannot explain all variation in the songs. There are a great number of influences on the pronunciation of singers which cannot be accounted for, and the mere recorded data does not allow for conclusions to be drawn. The fact that the variation can be explained by processes is out of the question, yet to account for all variation without considering different dialectal articulatory targets would be unsafe. It has to be said that the theory as well as the scope of this paper met its limits in this study, and even though there is phono-stylistic variation in pop songs, the question as to why it occurs remains yet to be answered.

References

Bannister, Matthew. 2006. "'Loaded': indie guitar rock, canonism, white masculinities". *Popular Music* 25/1, 77-95.

Banse, R. & K. Scherer. 1996. "Acoustic Profiles in Vocal Emotion Expression". Journal of Personality and Social Psychology 70/3, 614-636.

Bjarkman, Peter C.. 1986. "Natural Phonology and Strategies for Teaching English/ Spanish Pronunciation". In Bjarkman, Peter C. & Raskin, Victor (eds.). *The Real-World Linguist: Linguistic Applications in the 1980s*. Norwood: Ablex, 77-115.

Carlsson, Carl Johan. 2001. "The Way The Sing It: Englishness and Pronunciation in English Pop and Rock". *Moderna Språk* 95, 161-168.

Donegan, Patricia. 1978. *On the Natural Phonology of Vowels*. Ann Arbor: London University Microfilms International.

Donegan, Pantricia & Stampe, David. 1979. "The Study of Natural Phonology". In Dinnsen, Daniel A. (ed.). Current Approaches to Phonological Theory. Bloomington: Indiana University Press, 126-173.

Donegan, Pantricia & Stampe, David. 2008. "Hypotheses of Natural Phonology (I and II)". http://ifa.amu.edu.pl/plm/ (accessed on 17 October 2008).

Dressler, Wolfgang U.. 1982. "A Semiotic Model of Diachronic Process Phonology". In Lehmann, Winfried P.. & Malkiel Yakov (eds.).

Perspectives on Historical Linguistics. Amsterdam: John Benjamins Publishing Company, 93-131.

Dressler, Wolfgang U.. 1984. "Explaining Natural Phonology". In Ewen, Colin J. & Anderson, John M. (eds.). *Phonology Yearbook. 1*. Cambridge, CUP, 29-52.

Dressler, Wolfgang U.. 1996. "Principles of naturalness in phonology and across components". In Hurch, B. & R. A. Rhodes (eds.). *Natural phonology: the state of the art*. Berlin: Mouton de Gruyter, 41-52.

Dressler, Wolfgang U. & Moosmüller, Sylvia. 1991. "Phonetics and Phonology: A Sociopsycholinguistic Framework". *Phonetica* 48, 135-148.

Dressler, Wolfgang U. & Wodak, Ruth. 1982. "Sociophonological methods in the study of sociolinguistic variation in Viennese German". *Language In Society* 11/3, 339-370.

Fenk-Oczlon, Gertraud. 1989. "Geläufigkeit als Determinante von phonologischen Backgrounding-Prozessen". *Papiere zur Linguistik* 40/1, 91-104.

Foltin, Robert & Dressler, Wolfgang U.. 1997. "Natürliche Phonologie als interdisziplinäres Projekt". In Foltin, Robert & Dressler, Wolfgang U. (eds.). *Phonologie und Psychophysiologie*. Wien: Verlag der Österreichischen Akademie der Wissenschaften, 7-22.

Gibbon, Dafydd. 2007. "Formal is natural: Toward an Ecological Phonology". http://icphs2007.de/conference/Papers/1750/1750.pdf (accessed on 15 October 2008).

Hughes, Arthur, Peter Trudgill & Dominic Watt. 2005. *English accents and dialects. An introduction to Social and Regional Varieties of English in the British Isles*. (4th edition). London: Hodder Education.

Kärjä, Antti-Ville. 2006. "A prescribed alternative mainstream: popular music and canon formation". *Popular Music* 25/1, 3-19.

Ladefoged, Peter.1992. "The many interfaces between phonetics and phonology". In Dressler, W. U., Luschützky, H. C., Pfeiffer, O. E. & Rennison, J. R. (eds.). *Phonologica 1988. Preceedings of the 6th International Phonology Meeting*. Cambridge: CUP, 165-180.

Moore, Allan. 2002. "Authenticity as authentication". *Popular Music* 21/2, 209-223.

Moosmüller, Sylvia. 2007. *Vowels in Standard Austrian German. An Acoustic-Phonetic Analysis*. Habilitationsschrift an der Universität Wien.

Ohala, John J.. 2005. "Phonetic explanations for sound patterns: Implications for Grammars of Competence".
http://linguistics.berkeley.edu/PhonLab/users/ohala/papers laver-pdf.pdf. (accessed on 19 February 2009)

Pardo, Jennifer S.. 2006. "On phonetic convergence during conversational interaction". *Journal of the Acoustical Society of America* 119/4, 2382-2393.

Rhodes, Richard A.. "English reduced vowels and the nature of natural processes". In Hurch, B. & R. A. Rhodes (eds.). Natural phonology: the state of the art. Berlin: Mouton de Gruyter, 239-260.

Roach, Peter. 2000. *English Phonetics and Phonology. A practical course*. (3rd edition). Cambridge: CUP.

Scherer, Klaus R.. 1981a. "Speech and Emotional States". In Darby, John K. (ed). *Speech Evaluation in Psychiatry*. New York: Grune & Stratton, Inc., 171-188.

Scherer, Klaus R.. 1981b. "Vocal Indicators of Stress". In Darby, John K. (ed). *Speech Evaluation in Psychiatry*. New York: Grune & Stratton, Inc., 189-220.

Scherer, Klaus R.. 1986. "Vocal Affect Expression: A Review and a Model for Future Research". *Psychological Bulletin* 99/2: 143-165.

Stampe, David. 1969. "The Acquisition of Phonetic Representation". Papers from the Fifth regional meeting of the Chicago Linguistic Society, 443-454.

Stampe, David. 1979. *A dissertation on natural phonology*. Garland Publishing, Inc. : New York & London.

Stampe, David. 1984. "On Phonological Representation". In Dressler,W.U., Luschützky, H. C., Pfeiffer, O. E. & Rennison, J. R. (eds.). *Phonologica 1984. Preceedings of the 5th International Phonology Meeting*. Cambridge: CUP, 287-301.

Trudgill, Peter. 1997. "Acts of Conflicting Identity: The Sociolinguistics of British Pop-song pronunciation". In Nikolas Coupland & Adam Jaworski (eds.). *Sociolinguistics: A Reader.* New York: St Martin's Press, 251-265.

Wells, John Christopher. 1982. *Accents of English 2. The British Isles*. Cambridge: CUP.

Discography

The Coral, Jacqueline. Deltasonic Records 2002 Limited, 2007.

The Coral, Jacqueline (Live). Deltasonic Records 2002 Limited, 2007.

Dogs, She's Got A Reason. Island, 2005.

Good Shoes, Ice Age. Brille Records Ltd., 2007.

Good Shoes, Ice Age (Live). Brille Records Ltd., 2007.

The Jam, Town Called Malice. Polydor, 1982.

Lightspeed Champion, Tell Me What It's Worth. Domino Records, 2008.

Lightspeed Champion, Tell Me What It's Worth (Live). Domino Records, 2008.

Little Man Tate, Sexy In Latin. V2 Music, 2007.

Little Man Tate, Sexy In Latin (Live). V2 Music, 2007.

Ocean Colour Scene, If I Gave You My Heart. Island, 2001.

The Pipettes, Pull Shapes. Memphis Industries, 2006.

The Rifles, I Could Never Lie. Sixsevennine Recording Ltd., 2008.

Starsailor, Keep Us Together. EMI, 2005.

Travis, Happy. Independiente, 1997.

Paul Weller, All I Wanna Do (Is Be With You). V2 Music, 2008.

Paul Weller, All I Wanna Do (Is Be With You) (Live). V2 Music, 2008.

“All I Wanna Do (Is Be With You)”

I’m not out to convince you

ãm nɑt̪ ɑʊ tʰu kʰən'vĩn ʃuː

ãm nɒt̪ ɑʊ tʰuˑ kɒ̃n'vĩn ʃaː

You know I’m not that kind

ə nəʊ ãm nɒ θæ kʰãɪñd̪ʰ

jə nəʊ ãm nɑʔ θæʔ kãɪñ

I’m not here to begin ya

ãm nɑtʰ hɪə tʰʊ bə'g̃ĩn jaː

ãĩm nɑt̪ hɪe tʰʊ b̪ə'gĩn jaː

Don’t want you feeling used

d̪ə̃ʊ wə̃ʊ̃nt ʃuː 'fiːlĩn juːz̪d̪

d̪ə̃ʊ̃n wɒ̃n ʃuː 'fiːlĩŋ juːz̪

Or draw upon your mind

ə dɹɔːɹ̝ə'pɒ̃n jʊə mãĩnd̪ʰ

ɔː dɹɔːɹ̝ə'pɒ̃n jʊə mãĩnd̪ʰ

All I wanna do is be with you

ɔːɬ a 'wɒ̃nə duː ɪz b̪iː wɪð juː

ɔːl a 'wɒ̃nə d̪uː ɪz̪ biː wɪθ juː

I’m neither clever nor confused

ãĩm 'naɪðə 'kʰlevɐ nɔː kə̃n'fjuːz̪d̪

ãm 'naɪðə 'kʰlevə nɔː kə̃n'fjuːz̪

All I wanna do is be with you

ɑl a 'wɒ̃nə duː ɪz biː wɪð juː

ɑl a 'wɒ̃nə duː ɪz biː wɪð juː

I’m not looking to steal ya

ãĩm nɑt 'lʊkʰĩŋ tʰə stiːl jaː

ãĩm nɒʔ 'lʊkʰĩŋ tʰʊ stiːl jaː

I’m not looking to steal ya

ɑl aɪ 'wɒ̃nə duː ɪz biː wɪð juː

ɔːl a 'wɒ̃nə d̪uː ɪz̪ biː wɪð juː

“All I Wanna Do (Is Be With You)”

I bathe in dust
aɪ beɪð ĩn dʌstʰ
a beɪð ĩn dʌstʰ

I feel the most
aɪ fiːl ðə məʊstʰ
aɪ fiːɬ ðə mɒs

I twist and turn turn
aɪ tʰwɪst æ̃n tʰɜ̃ːn tʰɜ̃ːn
aɪ tʰwɪst æ̃n tʰɜ̃ːn tʰɜ̃ːn

I’m lost and found
ãĩm lɒst æ̃n fãʊ̃nd̥
ãĩm lɒst æ̃n fãʊ̃nd̥

In a moment
ĩn ə ˈmə̃ʊ̃mə̃ntʰ
n̥ ə mə̃ʊ̃mə̃ntʰ

In a single moment
ĩn ə ˈsĩŋgəɬ ˈmə̃ʊ̃mə̃ːn
ĩn ə sĩŋgl̩ mə̃ʊ̃mə̃ː

I’m not out to chain you
ãĩm nɑt̪ aʊ tʰə ʧ̃ʃẽĩn juː
ãĩm nɑt̪ ˈlʊkʰĩn tʰə ʧ̃ʃẽĩn ʃaː

Lead you forward from behind
liːʤ̃ʊ ˈfɑːwəd fɹʊ̃m bɪˈhãĩnd̥
liːʤ̃uː ˈfɔːwəd fɹʊ̃m bɪˈhãĩnd̥

I’m not here to pain you
ãĩm nɒt̪ hɪə tʰə pʰẽĩn juː
aɪ də̃ʊ̃ wɑ̃n tʰə tʰeɪk̚ dãːʊ̃

You know I’m not that kind
jə nəʊ ãm nɒt̪ θæ kʰãɪñ
jʊ nəʊ ãm nɑ θæ kʰãɪñd̥

All I wanna do is be with you
ɔːl a ˈwɒ̃nə duː ɪz b̥iː wɪθ juː
ɔːl a ˈwɒ̃nə d̥uː ɪz biː wɪð juː

All I wanna do is be with you
ɔːl a ˈwɒ̃nə d̥uː ɪz biː wɪθ juː
ɔːl a ˈwɒ̃nə d̥uː ɪz biː wɪð juː

Be with you
biː wɪð juː
biː wɪθ juː

Be with you
biː wɪð juː
b̥’iː wɪð juː

Be with you
biː wɪð juː
b̥’iː wɪθ juː

All I wanna do is be with you
ɔːl a ˈwɒ̃nə duː ɪz b̥iː wɪθ juː
ɔːl a ˈwɒ̃nə d̥uː ɪz biː wɪð juː

With you
wɪθ juː
wɪθ juː

Be with you
b̥iː wɪð juː
biː wɪð juː

“Ice Age”

I’m the one who makes mistakes

ãm ðə 'wõnəh hʉː meɪks mɪst'eɪks

ãm ðə 'wõnəɦ hʉː meɪks mɪsː'teɪks

I’m the one who can’t say sorry

ãm ðə wõn hʉː kʰɑ̃ːn seɪ 'sɒɹɛɪ

ãm ðə wõn hʉː kɑ̃ːnt seɪ 'sɒɹeː

I get bored far too easily

a ge̯ʔ bɔː faː tʰʊ 'iːzəlɪ

a geʔ bɔːʔ faː tʰʊ 'iːzəli

And I watch TV ‘cos I’m lazy

ænd̥ aɪ wɒʧ tʰiː'viː kʰɒz ãm 'leəzeə

æn a wɒʧ tʰiː'viː kɒz̥ ãm 'leɪzeə

I’m the one who says goodbye

ãm ðə wõn hʉː sez̥ ge̯ʊʔ'baɪ

ãm ðə 'wõnə huː sez gʊ'baɪ

I’m the one who doesn’t have the time

ãm ðə wõn hʉ̃ː dʌ̃zn̩ hæv ðə tʰãĩm

ãm ðə ʌ̃n hʉʊ dʌzn̩ hæv ðə tʰãĩm

I get bored far too easily

a ge̯ʔ bɔː faː tʰʊ 'iːzəlɪ

a ge̯ʔ bɔː faː tʰʊ 'iːzəli

And I watch TV ‘cos I’m lazy

ãnd aɪ wɒʧ tʰiː'viː kʰɒz ãm 'leɪz̥eə

ænd a wɒʃ tiː'viː kɒz ãm 'leɪz̥eə

I want to be the one who can dance without

aĭ wõn tʰʊɦ biː ðə wõn hʉː kʰə̃n dɑ̃ːns wɪ'vaʊʔ

a wõn tʰuː biː ðə wõn hʉː kʰə̃n dɑ̃ːns wɪ'ðaʊ

Getting drunk

ge'tʰɪn dɹʌ̃ŋk

ge̯'tʰɪn dɹɤ̃ŋk

I get bored far too easily

a geʔ bɔːə fɑː tʰʊ 'iːzəli

a geʔ bɔː fɑː tʰʊ 'iːzəli

And I keep repeating the same things

ænd̥ aɪ kʰiːp rə'pʰiːtʰɪn ðə sẽĩm θĩŋz

æn a kiːp rɪ'piːtʰɪŋ ðə sẽĩm θĩŋ æn

“Ice Age”

I’m the one who knows the truth

ãm ðə wɔ̃n hʉː nəʊ̯z ðə tɹuːθ

ãm ðə wɔ̃nə hʉː nəʊ̯z ðə tɹuːθ

I’m the one who pulls the news

ãm ðə wɔ̃n hʉː pʰʊɬ̯z ðə nuːz̯

ãm ðə wɔ̃n hʉː pʊɬ̯z ðə njuːz̯

I give up far too easily

a ɡ̟ɪv ɤp fɑː tʰə iːzəli

a ɡ̟ɪv ɤp faː tʰʊ iːzəli

And this just keeps on happening

ænd̯ ðɪs ʤəs kiːps ɔ̃n ˈhapənĩŋ

æn ðɪs ʤɤs kiːps ɔ̃n ˈhæpənĩn

I say Ice Age

aɪ seɪ aɪs eɪd̯ʒ

aɪ seɪ aɪs eɪd

Reading

ˈɹɪə̆d̯ĩŋ

ˈɹiə̆dĩŋ

Full frontal nudity and

fʊo ˈfɹɒ̃ntʰəɬ ˈnjʉːdɪtʰi ænd

fʊo ˈfɹɒ̃ntʰɐɬ ˈnjʉːdɪtʰi æn

I say Ice Age

aɪ seɪ aɪs eɪʤ

aɪ seɪ aɪs eɪd

Reading

ˈɹɪə̆d̯ĩŋ

ˈɻiːdĩŋ

Full frontal nudity

fʊo ˈfɹɒ̃ntʰəl ˈnjʉːdɪtʰi

fʊo ˈfɹɒ̃ntʰəl ˈnjʉːdɪtʰi æn

I’m the one who’s always right

ãm ðə wɔ̃n hʉːz ˈɒweɪz ɹaɪt̯

ãm ðə wɔ̃nə hʉːz̯ ˈɒweɪz ɹaɪ

I’m the one who can’t swallow my pride

ãm ðə wɔ̃n hʉː kʰɑ̃ːn ˈswɒləʊ ma pɹaɪd

ãm ðə wɔ̃n hʉː kɑ̃ːn ˈswɒləʊ ma pɹaɪd

We argue about the same things

wiː aːgjʉː ə̟ˈbaʊʔ ðə sɛ̃ĩm θĩŋz

wiː aːɡ̟jʉː əˈbaʊt̯ ðə sɛ̃ĩm θĩnz̯

We disagree on everything

wiː dɪsəˈgɹiː ɔ̃n ˈevɹɪθĩn

wiː dɪsɐˈgɹiː ɔ̃n ˈevɹəθĩn æn

I felt this way for a while

a felʔ θɪs weɪ fɔː ə waːɬ

a felʔ θɪs weɪ fɔː ə waɪɬ

“Ice Age”

And I’ve seen your face when you smile

ænd av sĩːn jʊə feɪs wen jʊ smaːɬ

æn av sĩːn jʊə feɪs wen jʊ smaːɬ

I’ve noticed the way you say I love you

aˑv ˈnəʊtʰɪs ðə weɪ jə seɪ a lɒv jʉː

av ˈnəʊtʰɪsd ð̩ weɪ jʊ seɪ a lɒv jʉː

We never dance the way we used to

wiː ˈnevə dɑ̃ːns ðə weɪ wɪ jʉːz̩ tʰʊ

wiː ˈnevə dɑ̃ːns ðə weɪ wɪ jʉːz̩ tʉː

I say Ice Age

aɪ seɪ aɪs eɪdʒ

a̰ɪ seɪ aɪs eɪdʒ

Reading

ˈɹɪə̆d̪ĩŋ

ˈɹ̥ːdĩn

Full frontal nudity

fʊo ˈfɹɒ̃ntʰəɬ ˈnjʉːdɪtʰi

fʊo ˈfɹɒ̃ntʰɐ ˈnjʉːdɪtʰi æn

I say Ice Age

aɪ seɪ aɪs eɪdʒ

aɪ seɪ aɪs eɪdʒ

Reading

ˈɹɪə̆d̪ĩŋ

ˈɹ̥ːd̪ĩːŋ

Full frontal nudity

fʊ ˈfɹɒ̃ntʰə ˈnjʉːdɪtʰi

fʊo ˈfɹɒ̃ntʰəl ˈnjʉːdɪtʰi

And I’ve been close to the

ænd aɪ bĩːŋ kʰɬəʊz̩ tʰʊ ðə

aː bĩŋ kɬəʊz̩ tʊ ðə

Edge and I’ve

ɛdʒ æn aɪ̯

ɛdʒ æn a

Been close to the

bĩŋ kʰɬəʊz̩ tʰʊ ðə

bĩŋ kləʊz̩ tʰʊ ðə

Edge and I’ve

ɛdʒ æn aˑ haˈhaːv

ɛdʒ æn a haˈhaɪ̯

Been close to the

bĩŋ kʰləʊz̩ tʰʊ ðə

bĩŋ kləʊz̩ tʊ ðə

Edge and I’ve

ɛdʒ æn av

ɛdʒ æn a

Been close to the

bĩŋ kʰləʊz̩ tʰʊ ðə

bɪŋ kləʊz̩ tʊ ðə

Edge

ɛdʒ

ɛdʒ

“Jacqueline”

You left before the rain came down
ʰʊ lef bɪˈfaː ðə ɻẽĩn kʰẽːm dãː
jə̯ lef biˈfaː ðɐ ɻẽːn kẽĩm dãː

December was the only sound
d̥ɪˈsẽmbə wʌz iː ˈõːnliː sãʊ̃
d̥ɪˈsẽmbə wʌz̥ iː ˈõːʊ̃nli sãʊ̃n

All the leaves fell to the ground
a̰ːl ðə liːfs feɬ tʉ ðə gɻãʊ̃n
äl ðə liːfs feɬ tʰuː ðə gɻãʊ̃n

When you went away
wẽn jʉ wẽntʰ əˈweɪ
wẽn juː wẽntʰ əˈweː

It’s a crying shame to see
ɪts ə ˈkɻäjə̃n ʃẽːm tʰə siː ə kʰɪs biˈkʌ̃m ə ˈmɛ̃ːmɒɻiː
ɪts ə kʰɻãĩn ʃẽĩm tʰə siː
ə kxɪs b̥iˈkʰɑ̃m ə ˈmɛ̃ːmɒɻiː

A kiss become a memory

Now I know just
na̰ʔa̰ noʊ t͡ʃʌs
naːʊ aɪ noː d̥ʒʌs

What you mean when you say
wʌt͡ʃuː mĩːn wẽn jə ˈseːə
ʍʌt͡ʃuː mĩːn ʍẽn juː ˈseːe

Way above where the north winds blow
weɪ əbɔˑv weə ðə nʌˑθ wĩnz bɬoː
weɪ əbʌv weə ðə nɔːθ wĩnz bləʊ

We will watch out afar to the valley below
wiː wɪl wʌt͡ʃ aʊ̯̰ əˈfaː tu ðə ˈvæliː biːˈloː
wiː wɪl wʌt͡ʃ aʊ̯̰ əˈfaː tə ðə ˈvæli biːˈloː

Oh Jacqueline I know
əʊ ˈd͡ʒækəlĩːn ãĩ nəʊ
əʊ ˈd͡ʒækəlĩːn ãĩ nəʊ

Oh Jacqueline I know
ʔəʊ ˈd̥ʒækxəlĩːn ãĩ nəʊ
oː ˈd͡ʒækəlĩːn ãĩ nəʊ

“Jacqueline”

A pattern of the ways it seems
ə ˈpʰæt̪ʰĩn ɑf ðə weɪz̪ ɪt sĩːmz̪
ə ˈpʰætə̃n ɒf ðə weɪ ɪt sĩːmz

Thats the way it has to be
ðæts ðə weː ɪt̪ hæz̪ tu biː
ðæts ðə weɪ ɪt hæz tə biː

Now I know just what you mean
nɑʊ ɑ̃ĩ nəʊ d̪ʒɐs ʍʌtʃʊ̃ː mĩːn
nɑʊ ɑ̃ĩ nəʊ dʒʌs wɒtʃʊ̃ː mĩːn

When you say
ʍẽn juː seːə
wẽn juː seɪ

Way above where the north winds blow
weː əˈbɔˑv weə ðə nɑː wĩnz̪ bloː
weɪ əbʌv weə ðə nɔːθ wĩnz blәʊ

We will watch out afar to the valley below
wiː wɪl wʌ̃tʃ ɑʊt̪ əˈfɑː tʰu ðə ˈvæliː biːˈloː
wiː wɪl wʌ̃tʃ ɑʊt əfɑː tə ðə væli biːləʊ

Oh Jacqueline I know
oː ˈd̪ʒækxəlĩːn ɑ̃ĩ nəʊ
əʊ ˈdʒækəlĩːn ɑ̃ĩ nəʊ

Oh Jacqueline I know
əʊ ˈd̪ʒækxəlĩːn ɑ̃ĩ nəʊ
əʊ ˈdʒækəlĩːn ɑ̃ĩ nəʊ

The Coven House where you now go
ðə ˈkʰəʊvə̃n hɑʊz̪ ʍea jũː nɑʊ goː
ðə ˈkəʊvə̃n hɑʊz wea jũː nɑʊ gəʊ

It stole our love don’t you know
ɪt stəʊl ɑʊ lɔˑv dõʊ̃ntʃ ũː nəʊ
ɑɪ stəʊl jʊə lʌv dəʊ̃ntʃ ũː nəʊ

Don’t you know
dõʊ̃ntʃ ũː nəʊ
dəʊ̃ntʃ ũː nəʊ

Oh Jacqueline don’t go
əʊ ˈd̪ʒækəlĩːn dəʊ̃nʔ gəʊəʊ
əʊ ˈdʒækəlĩːn dəʊ̃n gəʊ

Oh Jacqueline I know
ˈd̪ʒækʰəlĩːn ɑ̃ĩ nəʊ
əʊ ˈdʒækəlĩːn ɑ̃ĩ nəʊ

“Jacqueline”

Oh Jacqueline don’t go

əʊ ˈd̪ʒækʰəlĩːn dəʊ̃nʔ gəʊ

əʊ ˈd̪ʒækəlĩːn dəʊ̃n gəʊ

Oh Jacqueline I know

əʊ ˈd̪ʒækəlĩːn ãĩ nəʊ

əʊ ˈd̪ʒækəlĩːn ãĩ nəʊ

“Sexy In Latin”

Well our mum’s got talking in the village store

hʊəɬ aʊ mʊ̃mz gɒʔ ˈtɔːkɪ̃n ɪ̃n ðə ˈvɪlɪd̥ʒ ʃtɔː

həeɬ aʊ mʊ̃mz̥ gɒʔ ˈtʰɔːkɪ̃n ɪ̃n ðə ˈvɪɬɪʃ sɔː

I was three you were four

a̟ː wəz θɹiː juː wə fɔːə

aː wəz̥ θriː juː wəə fɔː

You looked lovely that’s for sure

juː ɬʊʔ ˈɬɤvɬi θas fə ʃʊə

juː ɬʊkd ˈɬʊvɬi θas fə ʃɤə

Just something about you

d͡ʒɤs ˈsʊ̃mθɪ̃ŋ əˈbaʊ jɜˑ

d̥ʒɤs ˈsʊ̃mθɪ̃n əˈbaʊ jəˑ

Together we went everywhere

təˈgeθə wiː wɛ̃n̪ ˈevɹɪwɛə

təˈgeθə wiː wɛ̃n ˈevrɪwɛa

You scratched my face I pulled your hair

jə ˈskɹæt̪ʃd ma feɪs a ˈpʊɬd ʒɐ hɛə

jɬ ˈskɹat̪ʃd ma feɪs aɪ ˈpʊɬd ʃə hɛə

You sent me tumbling down the stairs

jə sɛ̃n mɪ ˈtʰʊ̃mbᵊlɪ̃n dãʊ̃n ðə stɛə

ə̆ sɛ̃n mi ˈtʊ̃mbᵊlɪ̃n dãʊ̃n ðə stɛəz

Just something about you

d̥ʒʊs ˈsʊ̃mθɪ̃n əˈbaʊʔ jəˑ

d̥ʒɤs ˈsʊ̃mθɪ̃n əˈbaʊʔ jəˑ

You must have known that I want you

jɬ mʊst əv nəʊ̃n θaʔ ã wɒ̃n jə̃

jɬ mʊst əv nəʊ̃n θaʔ aɪ wɒ̃n jə

You must have known that I want you don’t you

jə̃ mɤst əv nəʊ̃n θaʔ ã wɒ̃n jɛ̆ d̥ʒʊ̃nt̪ jɤː

jə̆ mʊst əv nəʊ̃n θaʔ aɪ wɒ̃n jəˑ dɛ̆ʊ̃n jəˑ

Something’s happening

ˈsʊ̃mθɪ̃ŋz ˈhæpnɪ̃n

ˈsɤ̆mθɪ̃ŋz ˈhæpnɪ̃n

It’s written on the wall that you’re sexy in Latin

ɪs ˈɻɪʔɪ̃n ə̃n ðə wɔːɬ ðaʔ jə ˈseksɪ̃̽ɪ̃n ˈɬaʔɪ̃n

ɪs ˈɻɪʔn̩ ɒ̃n ðə wɔːɬ ðaʔ jə ˈseksɪ̃̽ɪ̃n ˈɬaʔɪ̃n

“Sexy In Latin”

This time come on
ðɪs ʧãĩm kʰʌ̃m ˈɒˈhɑ̃n
ðɪs tʰãĩm kʰõm ˈɒˈhõn

Everything’s changing now we can’t hold on
ˈevɹəθĩns ˈʧʃẽĩnʃn̩ n̩ wiː kʰãːn həʊɬd õn
ˈevɹəθĩns ˈʧʃẽĩnʃən n̩ wiː kãːnʧ hɐʊɬd õn

We stay out drinking in the park
wi steɪ aʊ ʤĩŋkĩn ĩn ðə pɑːk
wiː steɪ aʊʔ ʤ̊ʃĩŋkĩn ĩn ðə paːkʰ

I walk you home after dark
aɪ wɔːk jə hə̃ʊ̃m hᵊaːftə dɑːk
aɪd wɔːkʰ jə hə̃ʊ̃m aːftə daːkʰ

I didn’t mind it wasn’t far
a ˈdɪdn̩t mãĩnd ɪt ˈwa̰zn̩t faː
aˑ ˈdĩdn̩t mãĩnd ɪt ˈwa̰zn̩t p͡faː

There’s just something about you
ᵊs ʤ̊ʃʊs ˈsʊ̃mθĩn ˈəbaʊʔ jəˑ
s̪◡eəs ˈsʊ̃mθĩŋ ˈəbaʊʔ jəˑ

You went to university
jʉ wẽn tʰə jʉ̈ːnɪˈvɜːsɪtʰiː
jʉ wẽn tʰə jʉ̈ːnɪˈvɜːsɪtiː

You lost your virginity
juˑ ɬɒstʰ juə vɵˈdʒĩnətʰiː
juː ɬɒst juə vəˈʤ̊ʃĩnətiː

Saw more of him and less of me
sõm◡mɔːɹ əf ĩm ə̃n ɬes ɐ miː
sõm◡mɔːɹ əf hĩm ə̃n ɬes ɐ miː

There’s just something about you
ʤʊs ˈsʊ̃mθĩn əˈbaʊ jəˑ
ʤ̊ʃʊs ˈsʊ̃mθĩŋ əˈbaʊ jəˑ

You must have known that I want you
jɬ mɤst əv nə̃ʊ̃n θəʔ aɪ wõn jəˑ
jɬ mʊsʧ əv nə̃ʊ̃n θaʔ a wõn jə̰ˑ

“Sexy In Latin”

You must have known that I want you don’t you
jɬ mɤst əv nəʊ̃n θaʔ aɪ wɒ̃n jə dəʊ̃n jəˑ
jɬ mʊsʊ̯t əv nəʊ̃n θaʔ aɪ wɒ̃nt jəˑ d̥əʊ̃n jəˑ

Something’s happening
ˈsʊ̃mθɪ̃ŋz ˈhæpn̥ɪ̃n
ˈsʊ̃mθɪ̃ŋ ˈhapn̥ɪ̃n

It’s written on the wall that you’re sexy in Latin
ɪs ˈɹɪʔɪ̃n ɒ̃n ðə wɔːɬ ðaʔ jə ˈseksi◡ɪ̃n ˈlaʔɪ̃n
ɪs ˈɹɪʔn̥ ɒ̃n ðə wɔːɬ ðat ʃə ˈseksi◡ɪ̃n ˈɬaʔɪ̃n

This time come on
ðɪs tɐ̃ɪ̃m kʰʌ̃m ˈɒˈhɑ̃n
ðɪs tɐ̃ɪ̃m kʰʌ̃m ˈɑˈhɑ̃n

Everything’s changing now we can’t hold on
ˈevɹɪθɪ̃ns ˈʧẽĩnʃn̥ n̥ wiː kɐ̃ːnt həʊɬ ɒ̃n
ˈevɹɪθɪ̃ŋs ˈʧẽĩnʃn̥ n̥ wiː kʰɐ̃ːn həʊɬd ɒ̃n

Well we’re not friends anymore
weɬ wĩə nɒʔ fɹẽnz ˈæ̃nĩˈmɔː
weɬ wĩə nɒʔ fɹẽndz ˈæ̃nĩˈmɔː

I’m twenty-three you’re twenty-four
ɑ̃m ˈtʰwẽni θɹiː jə ˈtʰwẽni fɔː
ɑ̃m ˈtʰwẽni θɹiː jʊə ˈtʰwẽni f’ɔː

You’re still georgeous that’s for sure
jʊə stɪɬ ˈgɔːd͡ʒəs θas fə ʃɔː
jʊə sːɪɬ ˈgɔːd͡ʒəs θas fə ʃʊə

Still something about you
stɪɬ ˈsʊ̃mfɪ̃n əˈbaʊ jəˑ
stɪɬ ˈsʊ̃mp͡fɪ̃n əˈbaʊ jəˑ

Now we’ve done it everywhere
naʊ wɪ̯ d̥ʊ̃n ɪʔ ˈevɹɪwɛə
naʊ wɪ̯ d̥ʊ̃n ɪʔ ˈevɹɪwɛə

You scratch my back I pull your hair
jɬ skɹat͡ʃ ma bakʰ a pʊɬ jɐ hɛə
jə skɹat͡ʃ ma bakʰ a pʰʊɬ jʊə heə

“Sexy In Latin”

We’ve even done it on the stairs

wiːv 'iːvə̃n d̥ʊ̃n ɪʔ ɒ̃n ðə stɛəz

wɪv 'iːṽən dɒ̃n ɪʔ ɒ̃n ðə stɛəz

Still something about you

stɪɫ 'sʊ̃mfĩn 'əbaʊ jəː

stɪɫ 'sʊ̃mθĩn 'əbaʊ jəː

Something’s happening

'sʊ̃mfĩŋz 'hæpnĩn

'sʊ̃mfĩnz 'hæpnĩn

It’s written on the wall that you’re sexy in Latin

ɪs 'ɹɪʔĩn ə̃n ðə wɔːɫ ðaʔ jə 'seksi‿ĩn 'ɫaʔĩn

ɪs 'ɹɪʔn̩ ə̃n ðə wɔːɫ ðaʔ jə 'seksi‿ĩn 'ɫaʔĩn

This time come on

ðɪs t^{h}ãĩm k^{h}ʌ̃m 'ɒ'hɒ̃n

ðɪs sãĩm k^{h}ʌ̃m 'ɑ'hɒ̃n

Everything’s changing now we can’t hold on

'evɹɪθĩns 'ʧẽĩnʃn̩ ĩn wiː k^{h}ãːn həʊɫ ɒ̃n

'evɹəθĩns 'ʧẽĩnʃn̩ n̩ wiː k^{h}ãːnt həʊɫd ɒ̃n

Something’s happening

'sʊ̃mfĩnz 'hæpnĩn

'sʊ̃mfĩnz 'hæpnĩn

It’s written on the wall that you’re sexy in Latin

ɪs 'ɹɪʔĩn ə̃n ðə wɔːɫ ðaʔ jə 'seksi‿ĩn 'ɫaʔĩn

ɪs 'ɹɪʔn̩ ə̃n ðə wɔːɫ ðaʔ jə 'sek^{h}si‿ĩn 'ɫaʔĩn

This time come on

ðɪs tãĩm k^{h}ʌ̃m 'ɒ'hɒ̃n

ðɪs t̬ãĩm k^{h}ʌ̃m 'ɑ'hɒ̃n

Everything’s changing now we can’t hold on

'evɹɪθĩns 'ʧẽĩnʃn̩ ĩn wiː k^{h}ãːn həʊɫd ɒ̃n

'evɹɪθĩns 'ʧẽĩnʃn̩ n̩ wiː k^{h}ãːnt həʊɫd ɒ̃n

No we can’t hold on

noː wiː k^{h}ãːnt̬ həʊhəʊɫd ɒ̆.ɒ̃n

nəʊ wiː k^{h}ãːnt̬ həʊhəʊɫd ɒ̆̃.ɒ̃n

No we can’t hold on

noː wiː k^{h}ãːnt̬ həʊhəʊɫd ɒ̃.ɒ̃n

nəʊ wiː k^{h}ãːnt̬ həʊhəʊɫd ɒ̆̃.ɒ̃n

No we can’t hold on

noː wiː k^{h}ãːnt̬ həʊhəʊɫd ɒ̃.ɒ̃n

nəʊ wiː k^{h}ãːnt̬ həʊhəʊɫd ɒ̆̃.ɒ̃n

No we can’t hold

nəʊ wiː k^{h}ãːnt̬ həʊhəʊɫd ɒ.ɒ̃n

nəʊ wiː k^{h}ãːnt̬ həʊhəʊɫd ɒ̆̃.ɒ̃n

“Tell Me What It’s Worth”

Crack open the good times

kʰɹæk̚ ˈəʊpn̩ ðə gʊʔ tʰãɤ̆mz

kʰɹæk ˈəʊpʰən ðə gʊ tʰãĭmz̩

On the street corners busting rhymes

ɒ̃n ə stɹiː ˈkʰɔ̃ːnə ˈbʌstĩn ɹɑ̃ĭmz

ɒ̃n ə stɹiː ˈkʰɔ̃ːnəz̩ ˈbʌstĩn ɹãĭmz

But you fell between the lines

bʌˈʧuː feɬ bɪˈtʰwĩːn ðə lãĭnz̩

b̥ʌʔ juː feɬ bɪˈtʰwĩːn ðə laː.aaĭn

They all laugh become a joke

ðeɪ ɔːˌ̯læf bɪˈkʰɤ̃m ə ɖ̥ʃʒəʊkʰ

ðeɪ ɔː lɐf bɪˈkʰɤ̃m ə ɖ̥ʃəʊkʰ

Am I crazy baby let’s all hope

æm aɪ ˈkʰɹeɪz̩i ˈbeɪbi lets ɔːɬ hɐʊp

æm aɪ ˈkʰɹeɪzi ˈbeɪbi lets ɔːɬ həʊ

For narrow halls crunching drums

fə ˈnæɹəʊ hɔːɬz̩ ˈkʰɤ̆nʧĭn dɹɤ̃mz

fə ˈnæ.əʊ hɔːɬz̩ ˈkʰɤ̃nʧĭn dɹɤ̃mz̩

I got a sweet sugar but that’s all

a gɒt̬ ə swiːʔ ˈʃʊgə bʌ ðæts ɑːɬ

weɬ a gɑt̬ ə swiːt ˈʃʊgə bʌ ðæts ɔːɬ

Tell me what it’s worth

tʰeɬ miː wɒt̬ ɪts wɜː

tʰeɬ miː wɒt̬ ɪts wɜː

Tell me what it’s worth

tʰeɬ miː wɒt̬ ɪts wɜːɹʰ

tʰeɬ miː wɒt̬ ɪts wɜːʔ

So tell us that we’re spelling everything wrong

səʊ tʰel ʌs sæʔ wɪʔ ˈspelĩn ˈevɹəθĩŋ ɹɒ̃ŋ

tʰeɬ ʌs ðə wɪə ˈduːĩn ˈevɹəθĩŋ ɹɒ̃ŋ

Negroes turn a bluish grey when they’re dead

ˈniːgɹəʊz̩ tɜ̃ːn ə b̥luː.ˈɪʃ gɹeɪ wẽn ðeə ɖ̥’eəd

ˈniːgɹəʊz̩ tʰɜ̃ːn ə bluː.ˈɪʃ gɹeɪ wẽn ðeə ɖ̥’ed

“Tell Me What It’s Worth”

Well that’s funny cos I’ve just turned bright red

weɬ ðæts fʌ̃ni kʰəs av ɖ̥ʃʌs tʰə̃ːnd bɹaɪ̯t ɹed

weɬ ðæs fʌ̃ni a θĩŋk a tʰ̃ɜːnd ɹed

Kill kill kill

kʰɪɬ kʰɪɬ kʰɪɬ

kʰɪɬ kʰɪɬ kʰɪɬ

When everything starts to suck

wẽn 'evɹəθĩŋ stɑːts tʰə sʌkʰ

'evɹəθĩŋ stɑːts tʰʊ sʌkh

Drowning all your sins well I guess that’s bad luck

ɖ̥ʃãʊ̃nĩn ɔːɬ jʊə sĩnz weɬ a ges əs bæd lʌkʰ

ɖ̥ʃãʊ̃nĩn ɔːɬ jʊə sĩnz̥ a g̊es θæs bæd lʌk

Or the fact that your race is full of shit

ɔː ðə fæk θæt jʊə ɹeɪs ə fʊɬ ə ʃɪʔ

ɔː fæktʰ jʊə ɹeɪs ɪs fʊɬ ə ʃɪʔ

I got a sweet sugar but that’s it

a gɑ̯t ə swiːt 'ʃʊgə bʌʔ ðæs ɪʔ

weɬ a gɑ̯t swiːt 'ʃʊgə bʌ ðæts ɪʔ

Tell me what it’s worth

tʰeɬ miː wɒ̯t ɪts wɜː

tʰeɬ miː wɒ̯t ɪts wɜː

Tell me what it’s worth

tʰeɬ miː wɒ̯t ɪts wɜː

tʰeɬ miː wɒ̯t ɪts wɜː

Clean your blades and keep swinging

kʰ̃ĩˑn jə bleɪdz̥ ə̃n kʰiːp swĩŋ̃ĩn

kʰlĩːn jʊə bleɪdz ə̃n kʰiːp swĩŋ̃ĩn

Don’t stop till the red runs out

ɖ̥əʊ̃n stɒp tʰɪl ə ɹedə ɹ̆nz aʊ

dəʊ̃n stɒʔ tʰɪl ðə ɹedə ɹʌ̃nz aʊ

“Tell Me What It’s Worth”

Till no joy pours out of your mouth

tʰɪɬ nəʊ ʤʃɔɪ pɔːz̥ aʊt̬ əf jʊə maʊθ

tʰɪɬ nəʊ ʤʃɔɪ pɔːz aʊt̬ əf jʊə maʊ

Tell me what it’s worth

tʰeɬ miː wɒt̬ ɪts wɜːə

tʰeɬ miː wɒt̬ ɪts wɜː

Tell me what it’s worth

tʰeɬ miː wɒt̬ ɪts wɜː

tʰeɬ miː wɒt̬ ɪts wɜː

Tell me what it’s worth

tʰeɬ miː wɒt̬ ɪts wɜː

tʰeɬ miː wɒt̬ ɪts wɜː

Tell me what it’s worth

tʰeɬ miː wɒt̬ ɪts wɜː

tʰeɬ miː wɒt̬ ɪts wɜː

whoa-oh whoa-oh

ʍʊə.əʊ wʊə.əʊ

“All I Wanna Do (Is Be With You)” – Paul Weller

PROCESS	STUDIO	LIVE	DISCREPANCY
Voicing	5	5	0
Elision	10	15	+5
Monophthongisation	10	11	+1
Aspiration	21	18	-3
Insertion	0	2	+2
Devoicing	19	23	+4
Diphthongisation	1	0	-1

“Ice Age” – Good Shoes

PROCESS	STUDIO	LIVE	DISCREPANCY
Voicing	1	1	0
Elision	15	23	+8
Monophthongisation	23	28	+5
Aspiration	36	20	-16
Insertion	3	7	+4
Devoicing	29	23	-6
Diphthongisation	7	3	-4

“Jacqueline” – The Coral

PROCESS	STUDIO	LIVE	DISCREPANCY
Voicing	3	1	-2
Elision	10	10	0
Monophthongisation	13	7	-6
Aspiration	10	5	-5
Insertion	1	1	0
Devoicing	14	4	-10
Diphthongisation	0	0	0

“Sexy In Latin” – Little Man Tate

PROCESS	STUDIO	LIVE	DISCREPANCY
Voicing	6	8	+2
Elision	35	32	-3
Monophthongisation	12	8	-4
Aspiration	20	20	0
Insertion	14	15	+1
Devoicing	7	11	+4
Diphthongisation	2	2	0

“Tell Me What It’s Worth” – Lightspeed Champion

PROCESS	STUDIO	LIVE	DISCREPANCY
Voicing	13	11	-2
Elision	25	30	+5
Monophthongisation	4	5	+1
Aspiration	29	29	0
Insertion	1	4	+3
Devoicing	14	13	-1
Diphthongisation	2	0	-2

Total

PROCESS	STUDIO	LIVE	DISCREPANCY
Voicing	28	26	-2
Elision	95	110	+15
Monophthongisation	66	59	-7
Aspiration	116	92	-24
Insertion	19	27	+8
Devoicing	83	74	-9
Diphthongisation	12	5	-7

Zusammenfassung

Diese Arbeit beschäftigt sich mit phono-stilistischer Variation in britischen Pop Liedern der letzten Jahre. Das Ziel der Arbeit ist herauszufinden, welche Unterschiede es im Vergleich von Aufnahmen aus einem Tonstudio und Aufnahmen eines Live Auftrittes gibt. Als theoretischer Rahmen wird die Natürliche Phonologie gewählt, welche sich mit Prozessen beschäftigt, die die Variation in gesprochener Sprache beschreiben. Man unterscheided hier grundsätzlich zwischen zwei Prozessarten, *Stärkung* und *Schwächug*. Erstere, so wird argumentiert, erleichtert das Verständnis für die Zuhörenden, und Schwächung erleichtert die Produktion auf seiten der/des Sprechenden.

Als Datengrundlage sind fünf Lieder gewählt, werden phonetisch transkribiert und die Daten anschließend unter Gesichtspunkten der Natürlichen Phonologie analysiert. Ziel ist es herauszufinden in wie fern die Aussprache der Sänger von der jeweiligen Umgebung, also im Studio oder auf der Bühne beeinflusst wird. Basierend auf der Theorie werden Erwartungen an die Art der Prozesse und ihre Anwendung gestellt. So wird erwartet, dass in den Studio Aufnahmen Stärkungs-Prozesse öfter angewandt werden, und Schwächungs-Prozesse öfter in den Live Aufnahmen zu finden sind.

CURRICULUM VITÆ

Persönliche Daten

Name: Barbara Bucher

Geburtsdatum: 21. September 1985

Geburtsort: Innsbruck

Ausbildung

1995/96	Akademisches Gymnasium Innsbruck
1996 - 2003	Goethegymnasium Weimar, Abitur am 25.06.2003
seit 10/2003	Studium der Anglistik & Amerikanistik an der Universität Wien
01 - 05/2007	ERASMUS - Auslandssemester an der University of Manchester, U.K.