Investigating the Vowel Systems of Contact Languages using a Population of Artificial Agents

Stephen Dowell, M.A.

A thesis submitted in fulfilment of requirements for the degree of
Master of Science

to
Linguistics and English Language
School of Philosophy, Psychology and Language Sciences
University of Edinburgh

August 2006

Declaration

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgement is made in the text.

Stephen Dowell

Abstract

Contact languages are often described as having very simple phonological systems, with a small itinerary of vowels. This study analyses whether a computational architecture implemented to investigate the emergence of vowel systems (de Boer 2001) can be used to model these findings from the contact linguistics literature.

Simulations were run to generate emergent vowel spaces (much like those of de Boer 2001) which were then used to generate contact simulations.

Two kinds of contact simulations were explored. A "pidgin" simulation in which agents were taken from two different simulations, and a "creole" simulation featuring agents from two seperate simulations along with a number of "blank" agents (agents with no vowels in their vowel spaces).

Simulations in which vowels systems are generated from "scratch" provide similar results to those reported by de Boer (2001), but results from the contact simulations do not explain the patterns we encounter of real contact language vowel systems.

"Pidgin" contact systems were discovered to contain a high number of vowel clusters and the "creole" systems fewer. This contradicts the language literature.

Reasons for these findings are explained on the basis of shortcomings in the model and the complexity in modelling contact languages.

Acknowledgements

I would like to thank Kenny Smith for being a great supervisor, helping me with theoretical, coding and write-up issues. I would like to thank Simon Kirby for the suggestions and criticisms he provided early on. I would also like to thank all the people who kept me company in the lab, especially Cyp, Justin, Karola and Sebastian.

Contents

Chapter 5 Results 37

CHAPTER 1

Introduction

Though it seems clear that there are many species in the world capable of some form of communication, humans' ability to communicate with each other appears to be far more developed than the behaviour witnessed in other animals. Humans are able to discuss events and actions that happened in different places and at different times to the situation and moment of the conversation.

The human ability for language, unparallelled by any other species, provides us with a number of interesting questions about its origin, about why and how it came about. These questions are extremely difficult to answer as we cannot travel back in time to moments in our ancestral history to examine the communication system present at the time. All that we can do is examine language behaviour present in the world today and attempt to make educated inferences as to how language may have come about.

One interesting language phenomenon which could provide some insights is language contact. Language contact refers to the coming together of different languages in a population of people. The term "language contact" is used to refer to a large number of different sociological situations with differing results (see next chapter for an overview of the field).

Often considered the most extreme form of language contact, pidginisation describes the birth of a new language (pidgin language) which arises when a population containing speakers of different native languages is brought together and forced to create a communication system. Pidgins are most commonly associated with scenarios where a population of slaves (most probably from different places and hence with different native languages), all forced to work for some colonial masters (speaking their own language, generally illustrated as English, Spanish, Portuguese or Dutch in examples), have to learn the language of their masters as best they can.

The resulting languages, therefore, tend to have features in common with the languages of the colonial masters involved. This is normally observed in the vocabulary of the pidgin which tends to be heavily based on the vocabulary of the language spoken by the dominant speakers.

Due to this fact pidgin languages are often viewed very derogatorily. Winford (2003) provides an example of this, talking of a flyer from a West Sussex book seller describing pidgins and creoles as "vulgar and debased English" as well as a quote from the work of a self-proclaimed student of the creole language Gullah (a contact language spoken off the coast of South Carolina):

"Slovenly and careless of speech, these Gullahs seized upon the peasant English used by some of the early settlers and by the white servants of the wealthier colonists, wrapped their clumsy tongues about it as well as they could, and, enriched with certain expressive African words, it issued through their flat noses and thick lips as so workable of form of speech that it was gradually adopted by the other slaves and became in time the accepted Negro speech of the lower districts of South Carolina and Georgia." (Gonzales 1922, pp.17-18).

Much of the contempt aimed towards pidgin languages is not simply an excuse to make racist remarks, as shown in the the Gonzales quote featured above. However, there are many people who believed that pidgin languages are "corrupting" the lexifier language they stem from.

“The pidgin English as spoken in these days is about the most atrocious form of speech perhaps one could find in any corner of the globe. It is neither one thing nor the other. Consisting of a mixture of Somoan and Chinese here and there, with an occasional word of Malayan, it is conglomeration truly worthy of the Tower of Babel.” (Quote taken from Romaine 1988, pp.10-11).

Pidgins were, in the past, also not fully appreciated by linguists. Romaine (1988) describes them as once being “the neglected step-children of linguistics” saying “they were thought to be marginal, and not ’real’ full-fledged languages” (Romaine 1988, p.1).

Linguists have, however, more recently realised the importance that pidgins could have.

Extreme language contact situations, which lead to the birth of pidgins, stretch people’s language abilities a great deal, requiring a huge amount of linguistic creativity.

Pidgins and creoles (a creole is described as a pidgin which has native speakers) can teach us a vast amount about language, whether it be some insight into our ability to acquire language or, as mentioned above, some ideas on how language came about in the first place.

A linguist who has a huge amount of interest in this area is Bickerton. Bickerton (2001, 1981) believes that by examining the regularities and structures in different creole languages we can determine how the language learning apparatus we have works, and from this, learn how language first originated. Bickerton worked on HCE (Hawaiian Creole English), a creole in his definition of the term, which meant it was created out of a pidgin which did not exist for more than a generation and arose in a population where only 20% of the speakers spoke the dominant language with the other 80% speaking a number of different languages. Bickerton gives this definition believing these are the creoles that should be examined as they are an excellent example of a situation in which human linguistic capacity is stretched to its uttermost.

Having recorded immigrant speakers of HPE (Hawaiian Pidgin English) and native speakers of HCE, Bickerton compared their language attempting to gain insights into the creativeness of the human language apparatus. He found that the creole speakers had made a number of innovations to the pidgin language and provides two possible alternatives for how:

1. They are produced by general problem solving devices
2. They are the result of "innate faculties" genetically programmed

Bickerton leans very much towards the second alternative saying:

"If Hawaii were the only place where people had been faced with the problem of reconstructing human language, it would be impossible to decide between these alternatives. However, Hawaii is far from unique. There are a number of creole languages in other parts of the globe, but produced under very similar circumstances. /.../ If some general problem-solving device were at work, we would not expect that in every different circumstance it would reach the same set of conclusions. /../ However, if all creoles could be shown to exhibit an identity far beyond the scope of chance, this would constitute strong evidence that some genetic program common to all members of the species was decisively shaping the result." (Bickerton 1981, pp.41-42).

Bickerton does explore the trends we find in creole languages across the world (Bickerton 2001, 1981) but he only examines the trends found in the morphosyntax of these languages.

There do, however, also appear to be trends in the phonology of pidgins and creole languages (these will be described in the next chapter).

But do we really need to posit that an "innate language faculty" is responsible for these trends? Or can we explain them using a more straightforward approach?

This paper will explore whether a process of self-organisation can be used to explain why there tend to be a general pattern in the vowel systems generally found in pidgin and creole systems around the world.

Self-organisation describes a process in which complex behaviour can emerge, not from an external source governing the whole system, but from local interactions within the system.

An example commonly given for self-organisation is that of termites building their mound (Oudeyer 2006, Lindblom & Studdert-Kennedy 1984). The intricacy and complexity of the mound may suggest that the termites building the nest were following some sort of blueprint for its design. This is not the case, however. The termites are simply obeying some simple intrinsic rules such as "if I come across a lump of earth, pick it up and place it where the pheromone signal is strongest" (Oudeyer 2006, p.4).

This dissertation will attempt to answer the question of whether a self-organising computer model of vowel acquisition can be used to model the vowel systems of contact languages with the following layout.

In chapter 2 a brief introduction to contact languages will be given as well as some information on the phonology of pidgin and creole languages.

In chapter 3 some previous models of phonological self-organisation will be reviewed.

In chapter 4 an overview of the implementation used in my work will be given.

In chapter 5 the results of various simulations performed will be analysed.

And finally, in chapter 6 the results obtained will be discussed along with any implications they may have.

CHAPTER 2

Language Contact

2.1 Introduction to Language Contact

Language contact is an extensive area of research amongst linguists. The phrase "language contact" is used to refer to a wide variety of phenomena, from less extreme cases like language borrowing to the more extreme examples of pidgins and creoles.

When situations arise where speakers of different languages come together in the same place, one or both of the languages can become affected by the other, or a new language can result. This phenomena is known as "language contact".

Language contact is used to refer to a number of different situations. According to Winford (2003) these can normally be grouped into three categories: Contact situations leading to language maintenance, those leading to language shift and those causing a new language to be created. Sebba (1997) identifies six possible consequences that can arise from language contact: Vocabulary and Grammar "Borrowing", code-switching, language convergence, pidginisation, creolisation and language mixing. To give a brief overview of the different phenomena grouped together in the notion of language contact I am going to use the framework employed by Sebba. Winford does discuss all of the six consequences put forward by Sebba in his three categories but I believe separating the details as Sebba does will also help to clarify a few important technical terms relevant to language contact along the way.

2.1.1 Borrowing

The term "borrowing" refers to the process of adopting a foreign language item and using it in some host language. In time the item becomes concretised as part of the host language. Winford defines borrowing as happening in a situation of language maintenance, meaning that the native speakers of the language doing the borrowing preserve the language from generation to generation with only minor changes. Borrowing is generally associated with the borrowing of lexical items from another language, keeping the phonology and syntax relatively untouched. Borrowed words in English include: "juice" from French, "tomato" from Spanish, "coffee" from Turkish and so on.

Almost all (if not all) languages in the world have undergone some kind of language borrowing and can therefore be thought of as a contact language to some extent (McMahon 1994).

2.1.2 Code Switching

Code switching refers to a situation in a community where bilingual or multilingual speakers who share the same languages alternate between their languages within the course of a conversation.

Here is a fine example of some code switching in an utterance by a young Puerto Rican girl living in New York City switching between English and Spanish.

"Hey Lolita, but the Skylab, the Skylab *no se cayó pa(-ra) que se acabe el mundo*. It falls in pieces. *Si se cae completo*, yeah. The Skylab *es una cosa que (e-)stá rodeando el* moon taking pictures of it. *Tiene tubos en el medio*. It's like a rocket. It's like a rocket.

(Hey Lolita, but the Skylab, the Skylab("didn't fall for the world to end"). It falls in pieces. ("If it falls whole"), yeah. The Skylab ("is something that's going around the") moon taking pictures of it. ("It has tubes in the middle")[repeated]. It's like a rocket [repeated]. (Zentella 1997, p.117)

2.1.3 *Convergence*

Much like code switching language convergence happens in a community in which the vast majority of speakers are bilingual or multilingual. Convergence occurs, though, when their languages are adjusted and become more similar to each other. Sebba (1997) gives the example of a village in India called Kupwar, in which the inhabitants had been multilingual in three languages (Urdu, Marathi and Kannada) for roughly 400 years. In those years the structure of the Kupwar varieties of these three languages became so similar that a sentence from one language could be translated into another by simply replacing the words of one for the other making no alterations to the syntax.

2.1.4 *Pidginisation*

Pidginisation is considered to be a more extreme form of language contact and is described as being the creation of a new language.

Some doubts do rise as to whether pidgins should be considered "languages" or simply dialects of the language that contributes the majority of their lexicon (Romaine 1988). However, the majority of linguists (Winford 2003, Thomason 2001, Sebba 1997, McMahon 1994, for example) consider them to be language in their own right.

A pidgin is a language created to allow speakers of different languages to communicate with each other. These languages are based on the speakers' native languages but they are, however, very simple, with simple phonology and syntax. Pidgins have no native speakers, they are second languages for everyone who speaks them.

2.1.5 *Creolisation*

A creole emerges in a settled community where the general language of communication is a pidgin. The creole comes into being when children grow up surrounded by the pidgin. They acquire the pidgin, or something similar to it, as their native language. The pidgin is, therefore, "promoted" to a creole as it now has native speakers. Creole languages do tend to be more complex than pidgins with fixed phonology, more complex grammars and so on.

2.1.6 Mixed languages

Mixed language contact situations are another example of an extreme language contact situation resulting in the creation of a new language, much like pidginisation and creolisation. Mixed languages are, however, rather different to pidgins and creoles as they are not the result of a group of people who do not speak the same language creating a medium to be able to communicate with each other, they are created in communities of multilingual speakers who have no difficulty communicating with each other. Generally what happens is that, in a bilingual community, a hybrid language is made by using the syntactic structure of one of the two languages with the vocabulary of the other. There are not a vast amount of examples for this type of language contact, quite possibly the most famous case is that of Media Lengua in Ecuador where words from Spanish are used with the syntax of Quechua.

2.2 Pidgin and Creole Phonology

"The disaster that can lead to the emergence of a jargon (a very primitive contact system) and subsequently of a more stable pidgin generally involves the migration of a socially dominated group. This can be in the context of slavery or of contract labour in a colonial setting. Often trade carried out an unequal footing is involved. A group of people is forced by circumstances to develop a new communication system, to be used with foreigners who speak a different language."
(Appel & Muysken (1987), p.175)

The preceding quote gives a brief summary on the extreme language contact phenomenon of pidginisation. This phenomenon leads to the birth of a completely new language as members of a population with different native languages attempt to communicate with each other. This is quite different to another form of language contact leading to the birth of a new language, mixed language, as in this situation, the members of the population are multilingual and can understand each other. The quote beginning this section talks about dominant groups and a group of people who are forced to learn the language of this dominant group. I will be using the term “superstrate” for the dominant group and “substrate” for the dominated group(s). These terms are also used to refer to the languages of the groups, i.e. superstrate language and substrate language(s).

2.2.1 Pidgin phonology

The opinion which is commonly held about the phonological structures found in pidgin languages is that they are very simple. They are said to have unmarked consonants and vowels (Sebba (1997) describes this notion of pidgins not using “highly marked” sounds as the avoidance of “difficult” sounds, p.39) and are said to be composed of simple CV (consonant-vowel) syllables, with no consonant clusters and no consonant-final syllables (Romaine 1988).

It has been commented on that the size of the phonological systems of pidgins is greatly reduced when compared to the superstrate and substrate languages involved in its birth (Hall 1966, taken from (Romaine 1988)).

One theory states that pidgin phonology has a reduced number of phonemes due to the fact that when the superstrate and substrate languages come into contact, the only phonemes that will remain in the pidgin are those which are common to all languages (Winford 2003).

Sebba (1997) highlights the importance of the substrate speakers’ native language, suggesting that they will pronounce the words of the jargon (pidgin) in a way that comes most “natural” to them (most probably using sounds from their own native language). Sebba also makes a suggestion that the phonological systems we find in pidgins could be due in part to the superstrate speakers attempting to “simplify” their speech or attempting to mimic the pronunciation of the substrate group or groups.

Sebba also discusses how when a pidgin becomes stable, the effect of the substrate languages decreases and the pidgin begins to acquire its own phonological norms.

2.2.2 *Creole phonology*

Mühlhaüsler who is interested in the phonology of pidgins and creoles talks about the expansion stage in a pidgin's development. As a pidgin expands, three phonological changes tend to occur: an increase in phonological distinctions, the emergence of phonological rules and the stylistic use of previously unstylistic variants.

The first of theses leads to distinctions introduced such as distinctions between [s] and [t] or [f] and [p] which were not present earlier or the expansion of the vowel system from system making five contrast to one making seven. Mühlhaüsler claims that this has happened in the vowel systems of West African Pidgin English, Torres Strait Broken and Tok Pisin (Mühlhaüsler (1986)).

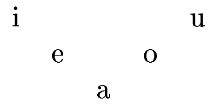


Figure 2.1: Five-vowel system shared by most pidgins (Mühlhaüsler 1986).

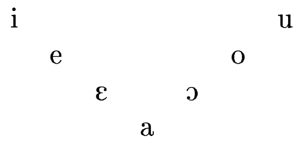


Figure 2.2: Seven-vowel system employed after expansion of pidgin (Mühlhaüsler 1986).

Winford (2003) posits that similar vowel systems to the ones portrayed in Figure 2.2 are present in the creoles of the world, saying that "practically all" have the five-vowel system with some having /ɛ/ and /ɔ/ also.

Bender (1987) (cited by Romaine 1988) does not seem to think that these two vowels (/ɛ/ and /ɔ/) are found in African Creoles presenting the following inventories for the vowels and consonants used:

consonants: /p, t, k, b, d, g, f, s, m, n, l-r, w, y/
vowels: /i, u, e, o, a/

Holm (2000, 1988) also sees the seven vowel system pictured above as a vowel system regularly found in contact languages. He marks it as the basic system of the Atlantic creoles. His reasoning for this is that it is the vowel pattern found in most of the West African languages which he claims many of the substrate speakers involved in the birth of the Atlantic creoles were likely to speak.

2.2.3 *Some oppositions*

Though many linguists do seem to believe that the world's pidgin and creole languages do seem to have many common aspects Thomason (2001) claims that these theories are, by no means, without exception offering a vast number of examples which violate the trends. As examples she discusses the consonant phoneme repertoire of Chinook Jargon which contains a number of marked consonant clusters, and the creoles Sarmaccan and Sango which both contain the double-articulated stop phonemes /kp/ and /gb/ as well as /p/, /k/, /b/ and /g/. Thomason also provides counterexamples for the popular claim that pidgin and creoles do not use tone distinctions, discussing a large number of contact languages that do, Ndyuka, Saramaccan and Sango being a select few of these. Winford (2003) also believes that there is a lot of diversity in the pidgins and creoles of the world saying: "This diversity and variation in pidgin lexicon and phonology contrast sharply with the uniformity of their reduced morphological and syntactic components"(Winford 2003, p.278).

2.2.4 *Overview of literature on pidgin and creole phonology*

The amount of work carried out into the phonology of pidgin and creole languages is very limited indeed. McMahon herself claims: "Relatively little attention has so far been paid to phonological creole universals" (McMahon 1994, p.261).

Though work on the subject has been sparse, the overall majority of what has been done seems to point to the fact that the phonology of pidgin and creole languages is quite simplistic with a limited number of phonemic contrasts and structural complexity. The vowel systems used, more specifically, are often basic five vowel systems when looking at pidgins which can expand to seven vowel systems when the pidgin becomes more stable and becomes “creolised” (starts to have children speaking it as a native language).

CHAPTER 3

Modelling the Self-Organisation of Speech

The universals of speech are most commonly explained as being the set of sounds available to humans, contained within an innate language device in the brain (Chomsky & Halle 1968). According to this theory when a child is growing up he or she will identify which of these universal sounds is specific to his or her language through the input language heard and set the correct parameters in the language centre of the brain.

Some researchers, however, have tried to show that the trends we find in the phonology of the world's languages do not need be explained using the existence of an in-built genetic ability for language. These researchers posit that the phonological universals we find can be explained, quite simply, as the outcome of a process of self-optimisation within a system. This could be a process in which, for example, the consonant system used in a certain language has taken the structure it has because, with time, it has been found to be the best system layout for which the consonants are easy to pronounce and at the same time distinct enough to make them distinguishable from one another.

3.0.5 Lindblom's models

Lindblom was the pioneer of research investigating the self-organisation of speech. In a paper with Liljencrants (Liljencrants & Lindblom 1972) he proposed that the sound systems of the world are the way they are due to the fact that they have optimised themselves so that their sounds can be easily pronounced but at the same time easily distinguished from each other. In this paper, Liljencrants and Lindblom looked at vowel spaces alone and posited the following equation to judge how good they were:

$$E = \sum_{i=0}^{n} \sum_{j=0}^{n} \frac{1}{r_{ij}^2},$$

where E is the energy of the of the vowel repertoire, n is the number of vowels in the system and r_{ij} is the perceptual distance between sound i and sound j measured by calculating the Euclidean distance between the two sounds' formant values. A lower energy function is considered better as it entails that vowels are more spread out from each other meaning that the possibility of them being confused for one another is decreased.

Liljencrants and Lindblom conducted simple computer simulations in which a certain number of points, representing vowels, were added to a vowel space (initially added in a circular pattern in the centre of the space). These points were then randomly changed, with the resulting change only being kept if the calculated energy of the repertoire was lower after the change.

Comparing the layout of the vowel spaces from their simulations with vowel spaces in real human languages showed a great deal of overlap.

This study performed by Liljencrants and Lindblom was very instrumental, as it showed that the trends in the structure of vowel systems around the world do not need to be explained using the idea of an innate language learning device. It showed that they could be explained, at a much more basic level, by simple principles, such as the need for vowels to be as far apart from each other as possible, acting on the system.

An important point to raise, however, is that the notion of vowel dispersion is not necessarily the same as vowel distinctness. Liljencrants & Lindblom (1972) refer to "maximal perceptual contrast" when discussing the findings of their model, meaning that the more dispersed a vowel system is (the lower its energy is examining the function above), the better its perceptual contrast. However, the idea of perceptual contrast is discussing the ability of the ears and relevant brain regions to distinguish sounds, whereas the representation of how distinct the vowels are is represented in acoustic space and is hence referring to how distinct the formant frequencies are. Though it would seem likely that these two representations are related to some degree it is by no means unquestionable.

In a different experiment (Lindblom & Studdert-Kennedy 1984, Lindblom 1986), Lindblom showed that a process of self-organisation can, in some sense, account for the phonological universals we find in language such as segments and features. In this model, phonetic signals had to be mapped to sets of “meanings” or “lexical elements”. The phonetic signals, portrayed as trajectories in an acoustic space, resembled stop-vowel syllables, however their structure did not presuppose a segmental analysis. They were “holistic” transitions running between an arbitrary point in the subspace of all possible stops and an arbitrary point in the subspace of all possible vowels.

The trajectories were optimised so as to be sufficiently distinct perceptually but also sufficiently easy to articulate. The results showed that some vowels were favoured, and that these vowels were paired with consonants in a complementary manner (i.e. if the place of articulation of the vowel was at the back of the vocal tract then it would generally be paired with a back vowel). These results matched with those found in real languages. They also found that many transitions shared the same start or end point, providing some evidence that the system was not just using “holistic” signals, but that they had been split up into two separate units. They therefore claimed that the results seemed to show that a certain amount of segmental or phonemic coding was taking place.

Both of the experiments discussed above, involving Lindblom, are very interesting as they allow us to see the phonological systems found in languages as having emerged in a rather simple way, through the outcome of a process in which they have been optimally adapted for distinctness and ease of articulation. However, as de Boer claims: “in reality one will rarely find human vowel systems that are completely optimally dispersed” (de Boer 2001, p.61).

Lindblom's models posit that the phonological systems we find in the world could be the result of self-organising processes taking place on the phonemes themselves. Other researchers have suggested that they may come about from the self-organisation of a population of speakers using them.

3.0.6 de Boer's model

A more recent study, using an agent-based model in which agents alter their vowel systems through imitation games, de Boer (de Boer 2001) has done a great amount of research on the self-organisation of speech. De Boer's model will be more clearly explained in the simulation section of the dissertation as the model used for this study is heavily based on de Boer's model, however, a basic outline will be given here. In de Boer's model, agents play language games, uttering vowels to be imitated or themselves imitating vowels that others have uttered. Agents carry a repertoire of vowels, and changes to their vowel spaces are made after a language game has ended, with different changes being made for different outcomes of the game. Changes are also made to the agents' vowel spaces which are not influenced by language games, such as ridding the space of unsuccessful vowels, merging vowels that are very similar, or randomly adding new vowels.

De Boer finds that in his computer simulations, vowel systems emerge that are shared among the agents and that these systems very closely match those found in real human languages.

De Boer's model is well received among phoneticians (a paper on his model was published in the Journal of Phonetics, (de Boer 2000)) as it, rather realistically, implements the agents' vowels in both articulatory and acoustic space.

It does receive some criticism among among linguists interested in the evolution of language, however, mainly for not explaining the origins of the processes which are employed in the model, with comments such as "De Boer's model does not explain, however, where the specific learning procedures come from." (Zuidema 2005, p.57) and "In particular, the question of why the agents try to imitate each other in de Boer's model is left unanswered (it is programmed in)."(Oudeyer 2006, p.61).

Despite these criticisms, however, de Boer's model is a good computational model which demonstrates interesting complex emergent behaviour from a relatively simple implementation structure.

3.0.7 Oudeyer's model

More recently still, another very interesting model of the self-organisation of speech has been proposed. Oudeyer (2006) provides a model composed of a population of agents able to produce and perceive sound, in which signals are represented in articulatory space alone. Agents in the model are equipped with two neural maps linked together, one represents a layer of perceptual neurons (which takes activation values from a model of the ear), the other a layer of motor neurons (which is linked to the perceptual layer with its output activations going to a model of the vocal tract).

Each agent in the simulation can perceive signals both by listening to another speaker, but also by vocalising signals themselves. As vocalising calls upon both the perceptual and articulatory map it causes the neurons activated in the two maps to be linked to each other. Listening to another agent causes activations in the perceptual map to be propagated to the articulatory map. Vocalising allows an agent to associate a given articulation with its perceptual counterpart and listening to another agent's vocalisation will increase the chance that the agent will use that vocalisation in future.

Oudeyer finds that the agents eventually settle on non-uniformly distributed vocalisations, meaning that agents have settled on a specific number of vocalisations that they use. The space of possible vocalisations chosen goes from being continuous to discrete. Oudeyer calls this reduction of signals to a small number of clusters "crystallisation". Extending his model, Oudeyer introduces a process analogous to that of cell death in the human brain, in which neurons that the agents do not use will eventually be removed. Oudeyer finds that the agents stabilise on a system in which all their neurons are being used enough and will, therefore, not be stripped away. The agents, therefore, seem to have agreed upon a shared set of phonemes, and as well as that, they seem to be able to recombine these phonemes in a phonotactically structured manner when producing signals.

I find Oudeyer's model very impressive and difficult to criticise, though there are, in my eyes, still issues regarding the realism of the model as there are with the other models in this section. Oudeyer emphasises the fact that his agents do not play the imitation game that de Boer uses in his model. He states about his agents: "Their interactions are not structured; there are no roles or coordination. In fact they do not have any social dispositions at all. They do not distinguish between their own vocalizations and those of others. They do not communicate. Here, 'communication' denotes the emission of a signal by an individual with the intention of transferring some information which will modify the state of another agent, which is not what happens here. In effect, the agents do not even know that there are other agents around them, so it would be difficult to say that they communicate." (Oudeyer 2006, p.86).

Clearly, Oudeyer is seeking an alternative to the, rather unrealistic, method of using language games (like imitation games) to model interactions among humans which could lead to phonological universals. I believe, however, that is quite unrealistic to posit that phonological norms could have come about without any form of communication among humans at all. When humans utter sounds they want other humans to engage with them, listen to them and, ultimately, be able to understand them. That is why I believe that, though very simplistic, the language games used in models examining the origins of language do characterise the human trait of wanting to interact with others rather well.

3.0.8 Some general points on the computational modelling of speech

The computational modelling of speech is an extremely important field. The theories they provide for the emergence of phonological norms have an advantage over the view that these norms exist because of an innate language faculty in our brains. Occam's razor encourages us to choose the most simple explanation for any phenomenon, not complicating it more than is necessary. I believe that the results from the models presented above provide us with more simplistic and more realistic justifications for the universals of speech.

CHAPTER 4

Simulation

The computer model used for this study is a mathematical model based on the work of de Boer (de Boer 2001). The model examines the self-organisation of vowels in a population of agents. The model is complicated enough to include some of the interesting features found in the way speech is produced and transferred but at the same time it is simple enough to understand without much effort, and simple enough for a computer to run in a short period of time.

Agents in the simulation have memory in the form of an articulatory store of the vowels in their vowel space, they have articulatory synthesiser allowing them to transform the vowels in their memory into an acoustic representation, and they have the ability to perceive an incoming acoustic signal and format the articulatory representation of vowels in their memory accordingly.

4.1 Agent Memory

Every agent in the simulation is equipped with the possibility to store an infinite number of vowels. Vowels are represented as a set of three real numbers between 0 and 1. These three numbers are values for the place, height and roundedness of the vowel's articulation. A value of 0 for the place of articulation of the vowel would indicate that the position of the tongue in the mouth is as far forward as it can possibly be and a value of 1 indicates that it is as far back as possible. For the height of articulation, a value of 0 indicates that the tongue is as high in the mouth as is possible and for a value of 1 the tongue is as low as possible. And for rounding, a value of 1 indicates that the lips are fully rounded for the articulation whereas a value of 0 indicates that the lips are completely spread out with no rounding. The layout of the values is (place, height, roundedness), so for the vowel [i], a front, high, unrounded vowel, the values would be (0, 1, 0), for [a] the values would be (0, 0, 0) as it is front, low and unrounded and for the back, high, rounded vowel [u], the values would be (1, 1, 1). Note that in the simulation these values can take on any real value between 0 and 1 and may look more like this (0.125, 0.985, 0.125) than the values given in the examples above.

These articulatory values of the vowels stored in the agent's memory can be interpreted into acoustic signals using the articulatory synthesiser.

4.2 The Articulatory Synthesiser

Acoustically, the vowels are represented as formant frequencies. In de Boer's model the first four formant frequencies are used, in this model only the first two are used. Using formant frequencies rather than the complete speech signal is perhaps less realistic as elements of the vowel such as volume, quality and so on are disregarded. However, using formant frequency values alone allows for a much faster simulation speed. De Boer claims "The first three or four are usually considered to be sufficient to represent a vowel". In this model only the first two are used. In discussing vowels in his book, entitled "The Sounds of the World's Languages", Ladefoged tends to only concentrate on the first two formant frequencies (Ladefoged & Maddieson 1996, chapter 9). For this reason and also because a four formant synthesiser would have been more time consuming to implement and run, only the first two formants were used.

The agent's internal articulatory representations of the vowels are synthesised into first and second formant frequency values using a set of mathematical equations. These mathematical equations (taken from de Boer 2001, p.44) are based on an interpolation

$$\begin{aligned} F_1 \quad = & \quad ((-392+392r)h^2+(596-668r)h+(-146+166r))p^2 \\ & + \; ((348-348r)h^2+(-494+606r)h+(141-175r))p \\ & + \; ((340-72r)h^2+(-796+108r)h+(708-38r)) \end{aligned}$$

$$\begin{aligned} F_2 \quad = & \quad ((-1200+1208r)h^2+(1320-1328r)h+(118-158r))p^2 \\ & + \; ((1864-1488r)h^2+(-2644+1510r)h+(-561+221r))p \\ & + \; ((-670+490r)h^2+(1355-697r)h+(1517-117r)) \end{aligned}$$

Figure 4.1: First and second formant articulatory synthesiser equations

function drawn from information gathered over a large number of vowels' formant frequencies found in Vallee's thesis (Vallée 1994, pp.162-164). De Boer describes the process as "an effective one" (p.45) saying that "the formant patterns that can be generated sound natural to human ears if synthesized". Although de Boer is referring to the synthesis of the first four formant frequencies generated from the equations, it does also suggest that the equations that convert from the articulatory positions of a vowel to its first and second formant frequencies are accurate. The first and second formant frequencies are given in Figure 1.1 where F_1 and F_2 are the first and second formant frequencies respectively, measured in Hz, *r* is the rounding value, *h* the height value and *p* the place of articulation value. These equations are extremely fast to process each requiring only 19 multiplications and 17 summations.

4.3 The Imitation Game

Agents in the simulation can update their vowel space by playing imitation games with each other. The term imitation game is based on the idea of language games put forward by Steels (1996, 1995). The language game, as described by Steels, is an interaction between at least two agents (a speaker and a hearer) who have a shared context. The interaction between the agents has a communicative dimension (e.g. the speaker may want to draw the hearer's attention towards a certain object), it has a linguistic dimension in that the speaker and hearer are negotiating about what signals should make up the language they share. Language games are a very simplistic way to model language interactions between humans. Though they are, perhaps, too simplistic in some respects, they do have certain interesting, quite realistic features. Agents involved in language games have internal representation of the language they have. The way they attempt to transfer this knowledge to other agent engaged in the language game with them is not by allowing them access to these mental representations but by using an external language system linked to the internal one. Agents in a population are able to converge upon an external language system with the help of some form of feedback given at the end of the language games. This seems to be in keeping with the way human language interactions work. People have an external form of language which they use to speak with one another. In doing this they most probably use a mental representation of language they have in their brains even though we do not understand how this actually works.

The imitation game is a language game in which the agents agree upon a shared language by attempting to imitate each other.

In this model the imitation game proceeds as follows: First an agent is chosen at random from the population of agents. This agent is designated as the initiator, chosen to start the imitation game. Another agent is also chosen at random (obviously this agent cannot be the same agent chosen to be the initiator). This agent is called the imitator and will engage in the imitation game with the initiator. The initiator then chooses one vowel at random from its memory. If the initiator has no vowels to choose from it will create a random vowel, add it to its memory and choose it for the game. A random vowel is simply created by generating three random numbers between 0 and 1 and setting each one to the three articulatory parameters (place, height, rounding). The initiator then synthesises the chosen vowel using its articulatory synthesiser. During the synthesis process a small amount of both articulatory and acoustic noise is added.

The articulatory noise is represented as a small random value added to the three articulatory parameters of place, height and rounding. The amount added is chosen from the uniform distribution within the range:

$$[\frac{-\Psi_{art}}{2}, \frac{\Psi_{art}}{2}],$$

where Ψ_{art} is the articulatory noise value, which defines how much articulatory noise is featured in the model (set at 0.1). The acoustic noise is represented in a similar way, adding random values to the first and second formant values. Due to the fact that human perception of pitch is logarithmic, however, the value for the acoustic noise added to the formant frequencies is adapted so that bigger noise values will be added to higher formant frequencies. The formant values with added noise are calculated with the following equation:

$$F_{noisy} = F_i(1 + v_i),$$

where F_{noisy} is the formant value with added noise measure in Hz, F_i is the formant value before noise is added (also measured in Hz), and v_i is a value taken from the uniform distribution:

$$[\frac{-\Psi_{ac}}{2}, \frac{\Psi_{ac}}{2}],$$

where Ψ_{ac} is the acoustic noise parameter (also set at 0.1).

Articulatory and acoustic noise are added to make the simulation more realistic. In communications among humans there may be articulatory slips or acoustic perturbations which disrupt the signal that is sent in a non-consistent manner. The addition of noise also ensures that no two vowels produced will ever be exactly the same.

The acoustic signal produced by the initiator, which has had noise added to it and is represented as the first two formant frequencies (F1 and F2), is perceived by the imitator agent. The imitator, having received the F1 and F2 values, then tries to find the closest vowel to the F1 and F2 values in its own vowel space, by taking each vowel in its inventory in turn, synthesising it, and calculating the distance from the two sets of formant frequency values. Although it may seem unrealistic, this is actually a rather realistic way to to model the way humans perceive sounds. According to de Boer "Humans tend to analyse speech sounds they hear in terms of sounds they already know" (p.47) and the use of the Euclidean distance is a good measure to classify how similar or different two vowels are perceptually. De Boer claims "the distance function assigns large distances to signals that are perceived to be very similar by humans, it is a good model of human perception.".

The distance is calculated using a straightforward Euclidean distance measure on the two sets of formant values. The equation used is:

$$D_{ac} = \sqrt{(F1_a - F1_b)^2 + \lambda(F2_a - F2_b)^2},$$

where D_{ac} is the acoustic distance, $F1_a$ is the first formant value of the first set of formants, $F1_b$ the first formant of the second set, $F2_a$ the second formant of the first set, $F2_b$ the second formant value of the second set and λ is the factor which weights the second formant value relative to the first formant value. This value is set at 0.3, the value used by de Boer which he based on experiments determining maximally dispersed vowel systems (Vallée 1994, Schwartz & Abry 1997) and independent data from production of vowels .

After determining which of the vowels in its store has the shortest distance to the formant frequencies perceived, the imitator then resynthesises this vowel and sends it back to the initiator.

The initiator perceives the vowel and then performs an operation just like the one performed by the imitator, going through all the vowels in its inventory and determining which vowel is closest to the signal. If the vowel it finds to be the closest to the signal it has heard is the same vowel as the one it initially chose to begin the game, then it will inform the imitator that the game has been successful, otherwise it will report a failed game (Lindblom & Lubker 1985).

Changes made according to the outcome of the imitation game are imposed solely on the imitator not the initiator.

A successful game has only one outcome: the imitator will move the vowel it chose slightly closer to formant frequencies it heard. This is done by creating six vowels, called neighbouring vowels, which differ by a small amount in one of the three possible dimensions (place, height, roundedness) positively or negatively (the amount they differ is set at 0.03). To explain using an example, the six neighbours of the vowel (0.5, 0.5, 0.5) are:

$$\begin{aligned}
(0.5 + 0.03, 0.5, 0.5) &= (0.53, 0.5, 0.5) \\
(0.5 - 0.03, 0.5, 0.5) &= (0.47, 0.5, 0.5) \\
(0.5, 0.5 + 0.03, 0.5) &= (0.5, 0.53, 0.5) \\
(0.5, 0.5 - 0.03, 0.5) &= (0.5, 0.47, 0.5) \\
(0.5, 0.5, 0.5 + 0.03) &= (0.5, 0.5, 0.53) \\
(0.5, 0.5, 0.5 - 0.03) &= (0.5, 0.5, 0.47)
\end{aligned}$$

The neighbours are all synthesised and the vowel whose formant frequencies are closest to the formant frequencies the initiator initially sent (measured using the Euclidean distance measure discussed above) is chosen to replace the initial vowel.

The procedure of moving the vowel closer to the acoustic signal heard, by picking the best neighbour, is performed to improve comprehensibility among agents. De Boer describes that the procedure is performed “in order to improve coherence in the population” (de Boer 2001, p.52) justifying it by saying “[c]hanging pronunciation in order to match others more closely is also necessary for children learning a language” (de Boer 2001, p.52).

Humans try out small variations on unfamiliar sounds in order to improve their pronunciation much like the agents in the model alter their vowels slightly by choosing a neighbouring vowel (de Boer 2001).

If a game is unsuccessful there are two possible actions that can be taken. These depend on how "successful" the vowel has been in the past.

Vowels in the agents' memory have counters for the amount of times they have been used and the amount of times they have been successful. Every time a vowel is used, by either the initiator or imitator, its use counter is increased by one. Every time a successful game is achieved, the vowel used (again for both initiator and imitator) has its success counter increased by one. The success ratio of a vowel is calculated by dividing the number of successes it has had by the number of uses it has had.

If the success ratio of the vowel the imitator used in the game is low (below 0.5), then the imitator will move the vowel it chose, when attempting to imitate the initiator's vowel, closer to the formant values it perceived. This is done by finding the six neighbours (as described above) of the vowel it used, finding the one with the shortest distance to the signal perceived, adding it to its vowel space with the use and success values of the original vowel, and then deleting the original vowel.

If the success ratio of the vowel used in the imitation game is high, then the imitator will add a new vowel to his space which closely resembles a vowel that, when synthesised, would produce formant values like the ones perceived. This feature of the model is used to avoid an agent moving a vowel which it has used successfully in the past. It is possible that the imitating agent needing to apply this "add similar vowel" feature has one vowel in an area of its vowel space where the initiating agent has two vowels. The agent, therefore, needs to add a new vowel to match the one it has heard rather than risk moving and disrupting the successful vowel it already has. The process of adding a new vowel is done by generating a new vowel in the middle of the articulatory space with values (0.5, 0.5, 0.5) and iteratively moving closer to the perceived formant values. This is done by finding the best neighbour of the generated vowel, then the best neighbour

```
1. F1 and F2 are the first two formants of the target vowel.
2. Create vowel tempVowel with values (0.5, 0.5, 0.5) for place, height and rounding.
3. Create neighbouring vowels of tempVowel and add them to list neighList along with
   tempVowel.
4. Check all vowels in neighList to make sure their values for place, height and round-
   ing all lie between 0 and 1. If they fail delete them from neighList.
5. Find vowel in neighList which, when synthesised, has the shortest distance to (F1,
   F2). Set as bestNeighbour.
6. If bestNeighbour does not equal tempVowel set tempVowel to be bestNeighbour and
   repeat steps 3-6. Otherwise quit.
```

Figure 4.2: Pseudo code explaining implementation of procedure used to add a vowel to the vowel space which is similar to the vowel heard

of the best neighbour and so on until the neighbour is no better than the vowel it is a neighbour of (see Figure 4.3 for pseudo-code which explains the algorithm).

4.4 Non-interaction Driven Changes

Aside from the changes that the agents make to their vowel inventories after the successful or unsuccessful outcome of an imitation game, the agents can also make other, non-interaction based changes. These changes are: adding new randomly placed vowels, deleting vowels with low success ratios, and merging vowels that are too close to each other. These three processes act upon the agents selected as initiator or imitator before the interaction between them takes place.

To allow for more vowels to be added to the space, agents in the simulation can add new random vowels. This happens with a low probability (set at 0.01) at each turn. When adding a random vowel, three random numbers are generated between 0 and 1 and are set as the three parameters of a vowel (place, height and rounding) which is then added to the agent's vowel inventory. De Boer implements this “in order to keep a pressure on the agent to utilize the acoustic space maximally”(de Boer 2001, p.54)

To ensure that the agent's vowel space is not littered with unsuccessful vowels, a clean-up procedure is used. With a certain probability (set at 0.1) of being executed at each turn, the method examines all the vowels present in the agent's inventory deleting those whose success ratio falls below a certain threshold provided they have been used a certain number of times (in this simulation they were set at 0.7 and 5 respectively, as they were in de Boer's model).

To avoid an agent having a cluster of vowels in close proximity which could cause confusion amongst each other, agents are able to merge two closely situated vowels into one. This is done for both articulatory and acoustic space at the beginning of every turn. All the vowels in an agent's space are compared to each other, their distances being measured. In acoustic space this is simply done by by subtracting the three parameter values of the second vowel chosen from the three parameter values of the first, squaring them, summing them together and then taking the square root of the sum. This is illustrated in the following equation:

$$D_{art} = \sqrt{(p_1 - p_2)^2 + (h_1 - h_2)^2 + (r_1 - r_2)^2},$$

where D_{art} is the articulatory distance, p_1, h_1 and r_1 are the place, height and rounding values of the first vowel in the comparison and p_2, h_2 and r_2 are the place, height and rounding values of the second vowel in the comparison. The distance below which two vowels will merge in this simulation is set 0.17 (the same setting which de Boer uses).

The distance below which two vowels will be merged in acoustic space is set to 2.2 times Ψ_{ac} (2.2 times the value for how much acoustic noise there is in the system) for this simulation. The acoustic distance is measured for the first and second formant in turn. The difference between the first formant value of the first vowel and the first formant value of the second vowel is calculated, ensuring that the result is positive. The result is then compared to the value obtained when the F1 value of the first vowel is multiplied by $2.2\Psi_{ac}$. The same process is repeated for the F2 values of the two vowels, and if both the comparisons result in the formant values of the second vowel being lower than or equal to the threshold value (2.2 times the distance that the value could be due to noise) then the vowels are merged.

Vowels are merged in both articulatory and acoustic space as it is possible that vowels with different articulatory positions may have the same acoustic representation and vice versa, that vowels with different articulatory positions may have similar acoustic representations. This is due to the fact that the articulatory synthesiser equations used are quadratic not linear and hence it is possible for them to produce the same result with different variable values.

When vowels are merged, both in articulatory and acoustic space, the better of the two vowels is kept (i.e. the one with the highest success score), the use and success values of the worse of the two vowels are added to those of the better one and the vowel is then removed from the space.

4.5 De Boer's Simulations

De Boer used the model described above to simulate the emergence of vowel systems. In these simulations, agents would start off with no vowels in their itinerary meaning that they would have to generate a vowel system from scratch. In order to replicate de Boer's findings and also so that data was available for subsequent follow-on experiments, 1,000 simulations were performed in which 20 agents started out with empty vowel spaces and a total of 5,000 imitation games were played (5,000 games were played to ensure a rather stable vowel system would be in place).

4.6 Language Contact Simulations

The model described above was then employed to investigate language contact scenarios. This was done by selecting two of the 1,000 simulations generated (as mentioned above), and combining them into one group who would then interact together. Two differing types of these language contact simulations were implemented.

4.6.1 “Pidginisation”

In the first scenario, “pidginisation”, ten out of the twenty agents in each of the two simulations coming into contact were chosen at random and grouped to form a new population of twenty agents. These twenty agents then played imitation games with each other (a total of 4,000 imitation games were played). This scenario was implemented to model the coming together of a population of adult speakers speaking different languages (see Section 2.1.4 for motivation).

4.6.2 “Creolisation”

In the second scenario, five randomly chosen agents from each of the two twenty agent simulations chosen for the contact situation were grouped into a population along with ten blank agents. In this simulation the ten blank agents and the ten agents with already existing vowel spaces (five from one initial simulation and five from another) also play 4,000 imitation games. This type of simulation was implemented to try and model the introduction of children into the population with no prior language model in place (see Section 2.1.5 for motivation).

4.6.3 General Details on Contact Simulations

To analyse the effect of contact within the model fully, contact simulations, of both types (“pidgin” and “creole”), were run on all the possible pairings of system size. The size of a system is measured by determining the average number of vowels per agent present in it. The size of all of the 1,000 simulations run was determined. The contact simulations were run on systems based on the size of their vowel system, ensuring that simulations were run for all the combinations of pairings. These pairings and their results will be discussed in the following chapter.

These contact simulations are very different to the simulations conducted by de Boer as there are agents in the simulation with already existing vowel spaces that are relatively well established. We would definitely expect these agents to have some kind of effect on the vowel system that emerges in the contact population.

CHAPTER 5

Results

In this section I will be looking at the results from a number of simulations performed on the model.

The results from the first series of simulations compare the behaviour of the model I built against the results de Boer reports from his simulation runs.

The subsequent simulations are designed to investigate whether de Boer's model could be extended to account for the claims made in the linguistics literature on the vowel systems of contact languages.

Analysing the behaviour of a simulation investigating the result of a contact situation between two vowel systems requires some means of measuring the difference or similarity between two systems.

I must note here that when I talk about "vowel systems" I am referring to the results of a whole simulation taking into account each vowel for every agent. If the repertoire of a single agent is being discussed then this will be made clear or the term "vowel space" will be used.

Analysing these "vowel systems" can, and will, be done by examining graphical representations of the system's vowels plotted in acoustic space. However, to attempt to produce an actual numerical figure for the difference or similarity of two systems a distance measure will also be used.

5.1 Measuring the difference between two vowel systems

The distance measure used to analyse how different/similar a system b is from a system a is simply an average distance for how close the closest vowel in system b is to every vowel in a.

For the distance measure to be calculated all the vowels for every agent in the two chosen simulations needs to be extracted and stored in a list.

The list is cleaned up in attempt to make the distance measure more accurate. This is done by removing any vowel in the list that may have been recently created and has not been involved in enough imitation games to be subjected to a vowel space clean-up. The number of imitations which a vowel needs to have been involved in to avoid being deleted is set to ten, in the hope that anything used a fewer number of times will have been thrown out already and anything used more times will have been on its way to being a successful vowel.

It is important to realise that the distance measure being explained is not symmetrical. When we talk about the distance between two systems, therefore, we also need to state the direction that we are looking at. The distance from system a to system b is not the same as the distance from system b to system a.

The distance from a system a to a system b is calculated as follows: Each vowel from system a is taken in turn. For each vowel the closest vowel from the other system (system b) in acoustic space is found, and the distance stored in a list of distances. The closest vowel is simply calculated by measuring the Euclidean distance between the formants of the chosen vowel and every vowel in the other system in turn, and choosing the vowel with the lowest distance value.

After this is done for every vowel in system a, the list containing all the shortest distance results is summed and then divided by the number of vowels present in the system.

This gives the average closest distance for all the vowels.

The equation for the distance measure is given below:

$$D(V_a, V_b) = \frac{\sum_{i=1}^{n_a} \left(\min_{j \epsilon \{1,2,\ldots,n_b\}} d(v_i, v_j) \right)}{n_a},$$

where $D(V_a, Vb)$ is the distance that vowel system a is from vowel system b, n_a and n_b are the total number of vowels in systems a and b respectively, and $d(v_i, v_j)$ is the distance between a single vowel i in simulation a and a single vowel j in simulation b.

5.2 Replicating de Boer's results

An important first step in exploring the model I implemented was to examine whether simulations in which agents started out with no vowels and then played imitation games with each other resulted in vowel systems emerging which were shared among the population and, to some degree, similar to the vowel systems encountered in real languages. These are the results that de Boer presents in his work. Therefore, to assess, firstly, whether my model was behaving correctly and, secondly, whether de Boer's results could be successfully replicated using the implementation he describes, the outcomes of these "emergence" simulations were examined.

The 20 agents involved in each simulation were found to successfully converge on a joint repertoire of vowels after having participated in the set 5,000 imitation games. This could be seen by the way that all the vowels for every agent within a simulation were grouped into clusters. These clusters were also found to be relatively distant from each other, meaning they were well dispersed in acoustic space.

A key method in the simulation was discovered to be the acoustic merge method which will merge two vowels into one if the Euclidean distance between them in acoustic space is below a certain value. This was found to be extremely important in determining the dispersion of the vowels in the system and hence the number of vowels that would emerge.

If the value was set too low, then vowels would not be merged unless they were very close in acoustic space meaning that clusters of vowels would form very close to each other and that each agent would have too many vowels in their itinerary. If the value was set too high, then vowels would merge together even if they were located far apart in acoustic space and so each agent would only have a very limited number of vowels in its space.

De Boer describes this minimum distance under which vowels should be merged as "two times the maximum distance over which the signal can be shifted owing to noise". This was interpreted to mean two times the acoustic noise value ($2\Psi_{ac}$).

This, unfortunately, produced a distribution of average vowels dissimilar to those reported by de Boer.

Figure 5.1 presents the distribution of the average number of vowels that de Boer reports his simulations have, calculated over 1000 runs.

When acoustic merge distance was set to $2\Psi_{ac}$ the distribution displayed in Figure 5.2 was obtained.

The distribution graphs exhibit peaks around whole numbers indicating that within a simulation, agents tend to have the same number of vowels. This fits well with de Boer's results. However, the number of resulting vowels appears to be too high with more simulations resulting in seven-vowel systems than six-vowel systems contrary to de Boer's which shows a far greater number of six-vowel systems.

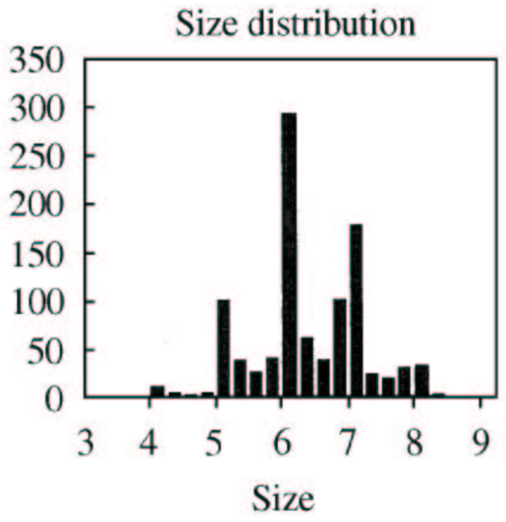


Figure 5.1: Histogram displaying de Boer's results for the average vowel space sizes when noise was set to 10%. The distribution is calculated over 1,000 simulations (de Boer 2000, p.15)

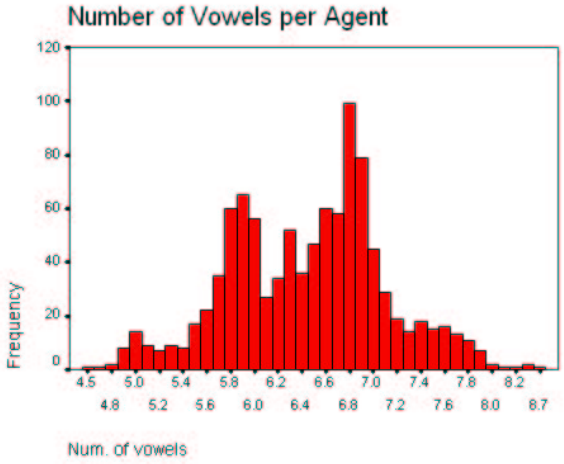


Figure 5.2: Histogram results for the average vowel space sizes when the model's merge distance was set to $2\Psi_{ac}$

The merge distance variable was, therefore, increased to $2.2\Psi_{ac}$ meaning that the range within which vowels would be merge was greater. The distribution illustrated in Figure 5.3 was then obtained.

We can see that, although it is not the same distribution obtained by de Boer, it is a closer match, with a higher number of six-vowel systems than any other number, with rela-

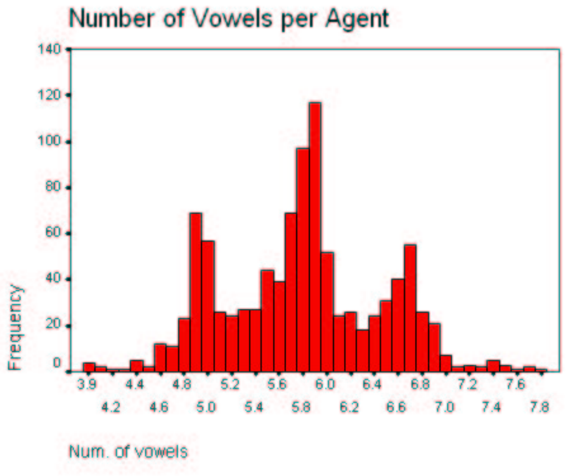


Figure 5.3: Histogram results for the average vowel space sizes when the model's merge distance was set to $2.2\Psi_{ac}$

tively high peaks at five and seven (though de Boer's distribution has a slightly higher peak at seven compared to five, whereas the opposite is true in this simulation), with a range from around 4 to 8.

The vowel systems that emerge also seem to bear certain similarities to vowel systems found by de Boer.

To plot the vowel systems in acoustic space I used the format of plotting F1 against F2-F1. This is a system used by Ladefoged (Ladefoged & Maddieson 1996) used to make the plot more suitable for a clear phonetic analysis of the vowels. According to Ladefoged "The acoustic representation corresponds more closely to the auditory phonetic description in terms of height and backness"(Ladefoged & Maddieson 1996, p.285).

A five-vowel system that emerged from one simulation is presented in Figure 5.4 which appears to have an itinerary of vowels similar to the system that de Boer reports as having an 88% chance of emerging in his simulations (de Boer 2001, p.95). The system is similar to an /i,e,a,o,u/ systems. The different coloured dots denote that the vowels belong to different agents (two dots of the same colour do not, however, necessarily belong to the same agent as there were not enough colours available to assign one to each agent).

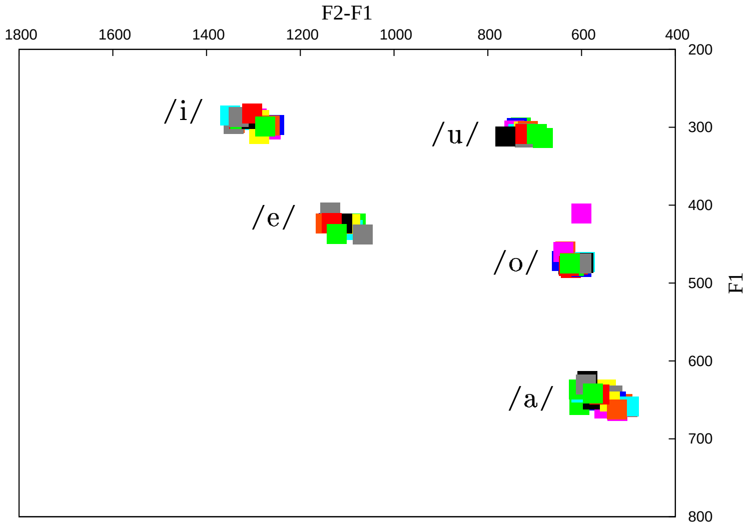


Figure 5.4: Example simulation of five-vowel system

An example of a seven-vowel system which emerged is presented in Figure 5.5. This is similar to the seven-vowel system of “Type B” that de Boer claims makes up 20% of the seven vowel systems he obtains (de Boer 2001, p.98).Its vowels are roughly /i,ɨ,u,ɛ,a,ɑ,ɔ/.

We can therefore conclude that the model does seem to work in the appropriate manner. The simulations show that agents converge on a shared system resembling those presented by de Boer.

5.3 Analysing the language contact simulations

The final systems of the two language contact scenarios were examined and compared. Two methods of examination were employed. These were the agents’ average number of vowels and the distance measure used to determine how different or similar one system is to another.

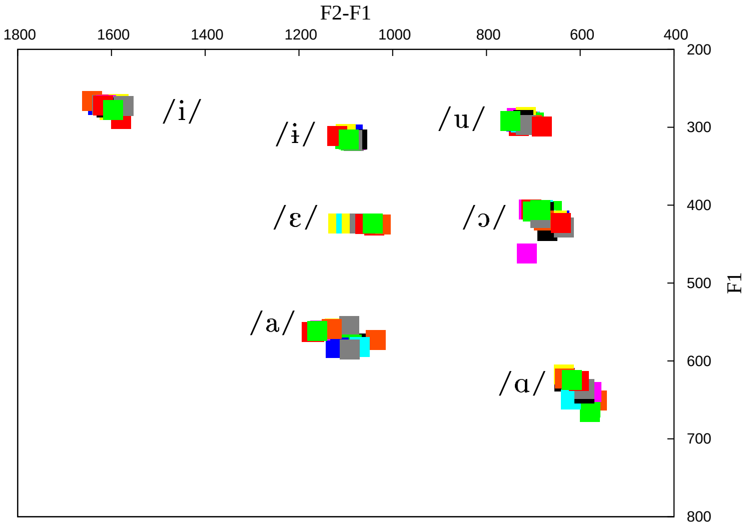


Figure 5.5: Example simulation of seven-vowel system

5.3.1 *Expectations*

If the computer model was to accurately portray what seems to be the trend in real pidgin and creole languages then we would expect the following:

1. A low average number of vowels per agent as an outcome of the "pidginisation" simulations (see Section 2.2.1).
2. If 1 is true then we may also expect the distance between the smaller system involved in the language contact situation (i.e. which has a smaller average number of vowels per agent) to be closer to the contact system. This expectation is based on the fact that certain linguists investigating contact languages believe that the vowel system of the resulting language is generally based on the vowel system of the substrate language(s) which is usually a smaller system (Sebba 1997). The distance between these two systems can be measured using the distance measure already discussed (see Section 5.1).

3. A higher number of vowels per agent in the results from the “creolisation” simulations where blank agents are introduced, compared to the “pidginisation” simulations (see Section 2.2.2).
4. We may also expect the difference between the bigger of the two systems coming into contact and the contact system to be smaller in the “creole” case than the “pidgin” case. This is based on the claim that when a pidgin is creolised it appears to gain vowels found in the superstrate language (Mühlhaüsler 1986, Sebba 1997), which is generally the system with the most vowels out of the languages coming into contact.

5.3.2 The selection of simulations

In order to investigate the effects of contact between systems of different sizes (see Section 4.6.3). The 1,000 simulations (see Figure 5.3 for the distribution) were sorted into the average number of vowels per agent they had.

Simulations were selected if their vowel size fell within 0.25 of a whole number. This was done as an attempt to select simulations where the agents seemed to have all mostly agreed on the same number of vowels.

As an example, if a simulation had an average number of vowels per agent of 5.15 it would be selected to feature in the set of vowels with 5 vowels. If, however, its vowel number was 5.4 it would not.

Number of Vowels	4	5	6	7	8
Number of simulations	7	205	350	88	3

Table 5.1: Table illustrating the number of simulations which had a suitable average vowel space size

Table 5.1 shows the number of simulations who have an average number of vowels per agent close to a whole number within the range of 4-8 vowels.

Due to the differing quantities of suitable simulations for each of the 5 groupings of vowel size, the number of contact simulations which could be generated by combining them varied. For pairings of vowel systems which would lead to only a few possible combinations being generated (below 2400), all the possible pairings of the initial simulations were explored. This was done for the pairings: "4,4", "4,5", "4,6", "4,7", "4,8", "5,8", "6,8", "7,8" and "8,8". For pairings of groups with a large number of simulations to choose from, a set number of contact simulations were performed choosing a pairing at random (this was set to 2000). This was done for the pairings: "5,5", "5,6", "5,7", "6,6", "6,7" and "7,7". The analysis for some of the contact simulations had to be deleted, however, due to computational limitations. Table 5.2 displays the number of contact simulations analysed for each pairing.

	4	**5**	**6**	**7**	**8**
4	21	1398	2385	592	16
5	-	1949	1956	1953	594
6	-	-	1959	1960	1029
7	-	-	-	1942	254
8	-	-	-	-	3

Table 5.2: Table showing the number of contact simulations analysed for each pairing of vowel system sizes (half the table is blank as pairing 6 with 4 is the same as pairing 4 with 6 for example)

5.3.3 Contact results

Examination of the results displays figures that completely contradict the expectations set out in Section 5.3.1.

The "pidginisation" simulations results in a high number of vowels per agent, with the number for the "creolisation" case being a considerable amount lower rather than higher.

In 5.6 we can see the mean values of the average number of vowels per agent for the two types of contact simulation for all the pairings featuring a five vowel system. This is representative of all the possible pairings which all follow a similar pattern. We can clearly see that the values for the "pidgin" case are consistently higher than those of

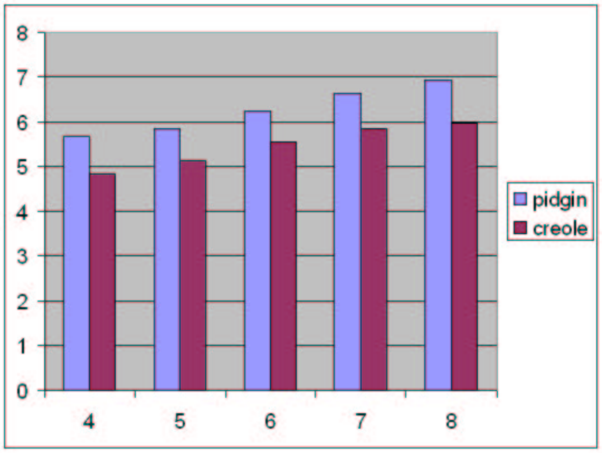


Figure 5.6: Bar chart showing the resulting average number of vowels per agent in the two contact simulations (“pidgin” and “creole”) when a five-vowel simulation is paired with simulations with differing numbers of vowels

the “creole”. It is clear from the graph above that the “contact” languages are not just making use of the vowel systems of the two languages that are involved in the contact simulation but also adding new vowels. This can be seen by looking at the average number of vowels per agent that emerges in the “pidgin” contact simulation when a four-vowel system comes into contact with a five-vowel systems or when two five-vowel systems come into contact. In both situations the average number of vowels per agent in the resulting “pidgin” contact simulation is close to six.

5.3.3.1 *Does the smaller system have any influence on the “pidgin” simulation?*

One question which arises from the outcomes witnessed is whether the smaller of the two systems involved in the language contact simulation has any affect whatsoever on the resulting “pidgin” system.

To check this question, the difference measure (discussed in Section 5.1) was employed to measure whether the distance between the smaller system (s) and the contact system (c) was smaller than the distance between the smaller system (s) and the bigger system (b).

Simulations s and c are completely unrelated. Their vowel systems are the outcome of completely different simulations. We would, therefore expect their vowel systems to be quite different and hence the distance between them to be quite large. Simulations s and c, however, are related as a number of agents, namely ten (see Section 4.6.1), are taken from simulation s, after it has finished running, and chosen to make up simulation c. We would, therefore, expect the distance between system s and system c to be closer than that between s and b.

Since the difference measure is not symmetrical we must investigate whether the distance from the small system to the contact system ($D(s, c)$) is smaller than the distance from the small system to the big system ($D(s, b)$), and also whether the distance from the contact system to the small system ($D(c, s)$) is smaller than the big system to the small one ($D(b, s)$). Figure 5.7 illustrates all the possible distances in a contact simulation.

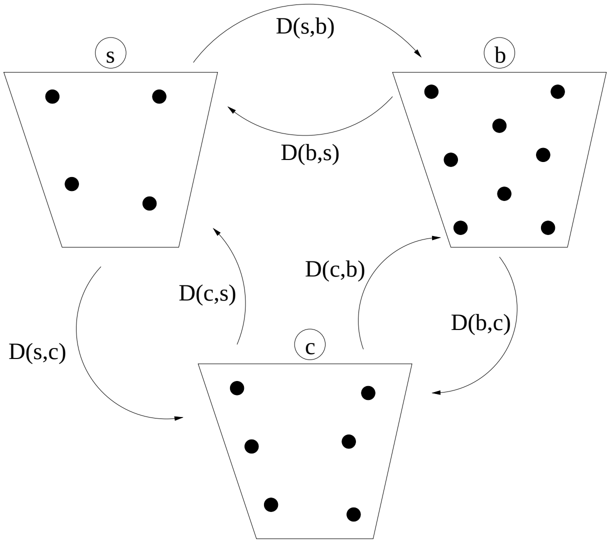


Figure 5.7: Diagram illustrating the names for all possible distance measures in a contact situation in which a smaller system (s) and a bigger system (b) are involved in a contact simulation resulting in a contact system (c).

We can observe from the results shown in Figure 5.8 that the smaller system is, quite clearly, having some effect on the outcome of the “pidgin” system as the distance from the smaller system to the contact one is consistently smaller than the distance from the smaller system to the bigger system.

The same is also true for the reverse of this distance. Figure 5.9 shows that the distance from the “pidgin” system to the smaller system is consistently smaller than the distance from the bigger system to the smaller system ($D(c, s)$ is consistently smaller than $D(b, s)$).

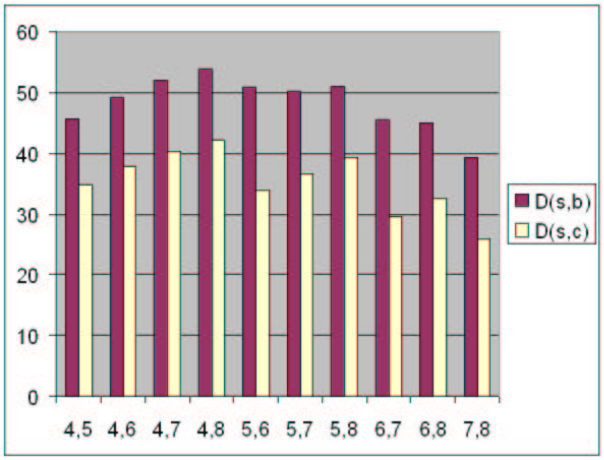


Figure 5.8: Bar chart comparing distances $D(s,b)$ (purple) and $D(s,c)$ (beige) in a "pidgin" contact simulation.

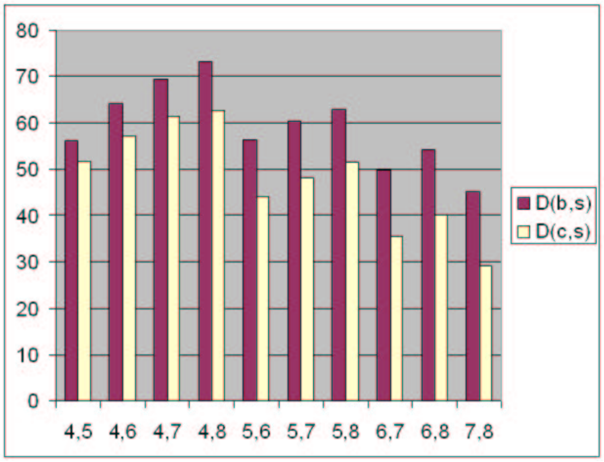


Figure 5.9: Bar chart comparing distances $D(b,s)$ (purple) and $D(c,s)$ (beige) in a "pidgin" contact simulation.

5.3.3.2 *Does the bigger system dictate the structure of the "pidgin" vowel system?'*

We can, therefore see that the smaller system is having some effect on the contact system, but how much of an effect is the bigger system having?

Do the "pidgin" simulations have a high average number of vowels per agent because vowels which the bigger system has, and the smaller system lacks anything similar to, are adopted in the system?

If this is the case we would expect the difference between the contact system and the bigger system to be relatively small. To assess this claim the distance between the bigger system and the contact system is compared to the distance between the bigger system and the smaller system.

Figures 5.10 and 5.11 show the results of this test, with Figure 5.10 showing the results of the difference measure applied forwards and Figure 5.11 in reverse (see Section 5.3.3.1 for an explanation of the two directions).

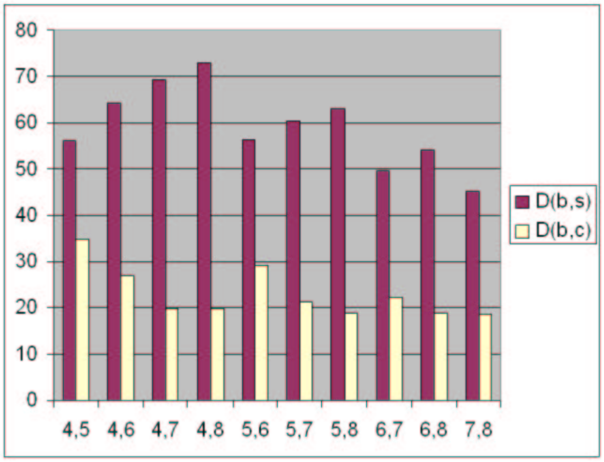


Figure 5.10: Bar chart comparing distances $D(b, s)$ (purple) and $D(b, c)$ (beige) in a "pidgin" contact simulation.

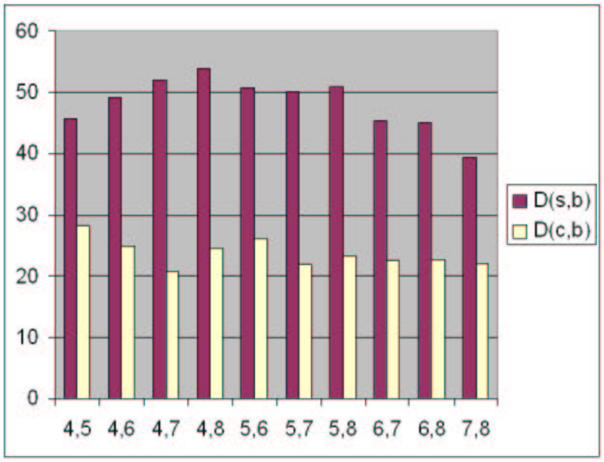


Figure 5.11: Bar chart comparing distances $D(s, b)$ (purple) and $D(c, b)$ (beige) in a "pidgin" contact simulation.

The results do seem to suggest that the contact system and the big system involved in it are similar, as the distance between them is relatively low especially when compared to the distance between the big system and the small system.

This would, therefore, seem to suggest that the bigger system involved in the contact simulation is contributing a great deal to the contact system.

Analysing the contact systems graphically also seems to point to the same conclusion.

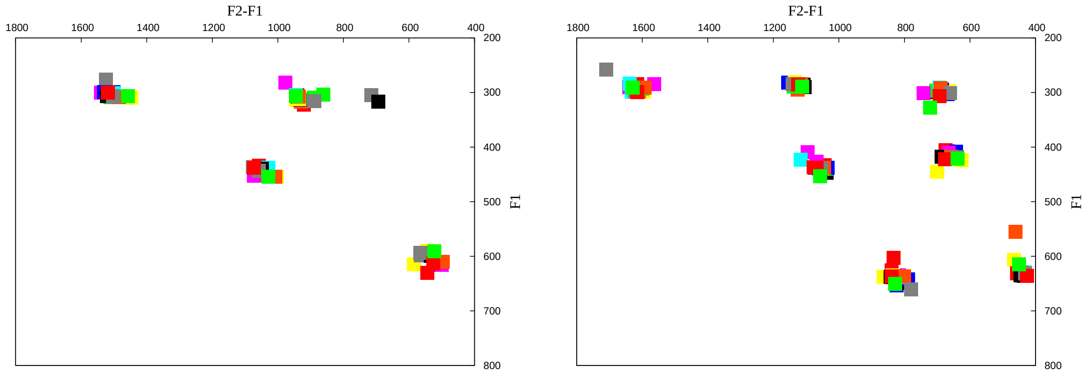
Figure 5.12 provides an example of the contact system that emerged when a four-vowel and a seven-vowel system (both chosen at random) were paired in a contact simulation. It, quite clearly, shows that the contact vowel system is very similar to the vowel system of the bigger system that was involved the contact simulation. Both the bigger system and contact system have seven vowel clusters all located in very similar areas. We must note, however, that the four clusters of vowels in the four-vowel system do seem to have similar correlates in the bigger system. Therefore, the fact that the contact system is very similar to the bigger system does not necessarily mean that it dissimilar from the smaller system. It seems to have adopted intermediary vowels for each of the four vowels in the bigger and smaller systems which are in similar areas of acoustic space, with the remaining three vowels being adopted from the bigger system.

The results from the analyses performed to determine the effect that the bigger system has on the contact system show that the effect of the bigger system is very important.

This is because if the large system has vowel cluster with no corresponding proximal clusters in the small system they will be included in the contact system without much change to them being made, as the smaller system will not influence their position a great deal.

5.3.3.3 *Do agents in the contact simulations continue to add vowels?*

We have seen that when a small and a big system come into contact, vowels from the big system which do not feature in the small system seem to be added to the contact system giving the contact system a similar layout to the bigger system. But what happens when two small systems are paired together? Figure 5.6 shows that on average when two five-vowel systems are paired together to form the contact simulation, the "pidgin" contact system will have an average number of vowels somewhere approaching six. How are



smaller system

bigger system

contact system

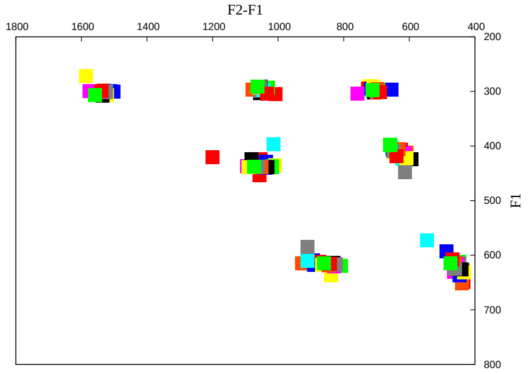


Figure 5.12: Figure displaying the resulting contact system when a contact simulation is performed on a 4-vowel and 7-vowel system.

these resulting six-vowel systems arranged? Do the basic systems making them up have clusters in different places to each other allowing a contact system to emerge using a set of six of these already existing vowels? Or do the agents introduce new vowels into the system not present in either of the two systems paired for the simulation?

Figure 5.13 examines the resulting system when two 4-vowel systems are paired to create a contact simulation. The contact system has, quite clearly, acquired vowel clusters at areas in the acoustic space where neither of the two initial systems had any. The clusters (circled in red in Figure 5.13) have, most probably, been added to the simulation during the contact simulation. This is due to the fact that agents are able to add a vowel at a random location in articulatory space with a certain low probability at the beginning of each game (see section 4.4). Due to the fact that the vowel systems coming into contact have a rather sparse set of vowels, the process of adding a new vowel has the chance to be

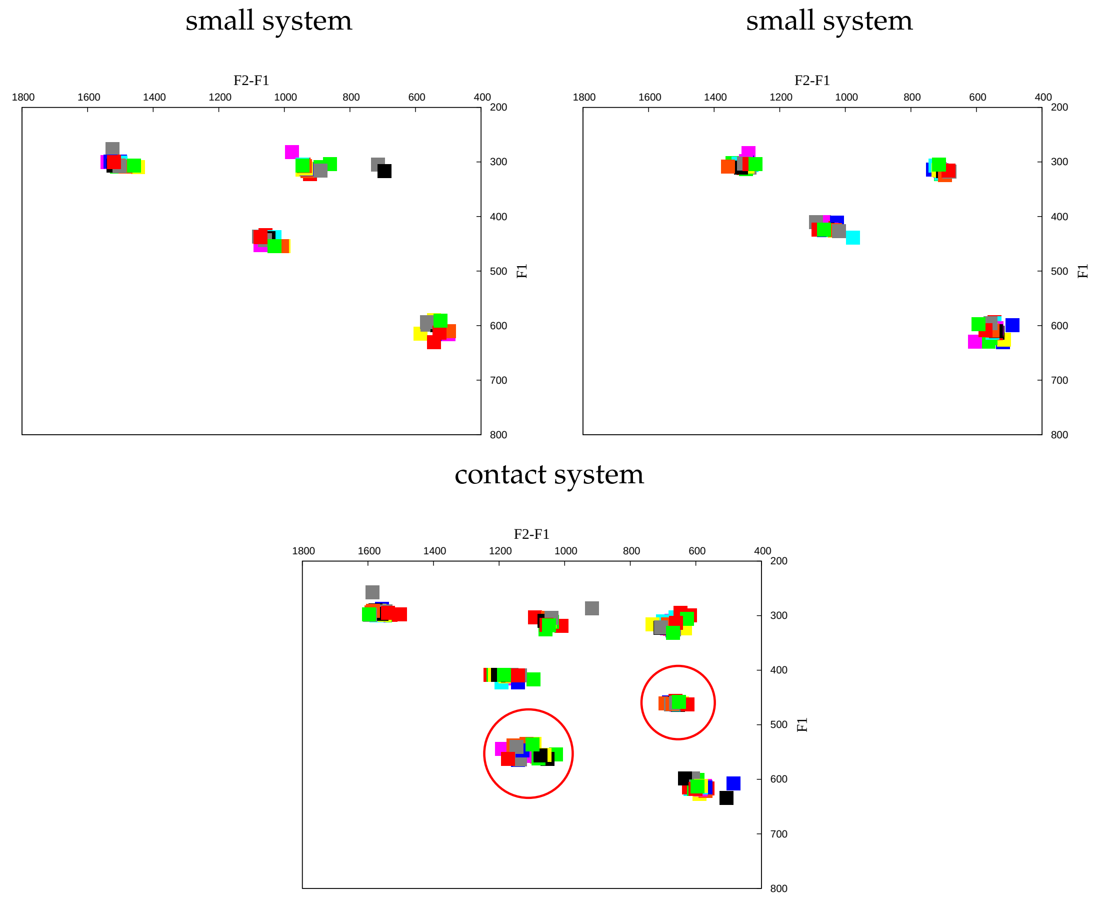


Figure 5.13: Figure displaying the resulting contact system when a contact simulation is performed on two 4-vowel systems. Circled in red are the vowel clusters which do not appear to have any correlates in the two systems involved in the contact simulation.

successful as there are areas of the space where vowels will be distinct enough from other vowels meaning they will not be merged to another. This feature of the model could have been excluded for the contact simulations meaning that agents in the contact simulations could not have added any new vowels to the contact vowel system. However, I wanted to keep the model as consistent as possible for both types of simulation (emergent and contact), exploring the behaviour of the model, that produces emergent vowel systems, when used in a contact situation.

CHAPTER 6

Discussion

In pidgin languages we generally find a very simple vowel system, normally with only five vowels (/i,e,a,o,u/). When the pidgin stabilises and young children begin speaking it as a native language the pidgin is "creolised" becoming a creole. Creoles are believed to have more complicated phonologies than pidgins often adopting a seven-vowel system.

This study attempted to simulate very simplistic models of language contact situations in groups of artificial agents.

The results achieved, however, do not match with the contact linguistics literature, with the results of "pidginisation" simulations showing a majority of vowel systems with a high number of vowels in them and the "creolisation" simulations producing vowel systems with, on average, fewer vowels than the "pidginsation" simulations.

The results obtained present us with a number of interesting questions as to why the model did not produce vowel spaces which we would expect to see in real world pidgin and creole languages.

6.1 Justifications for the results

6.1.1 *The model is too simple*

One possible explanation for the results achieved being so diverse to what we find in the real world could be due to the design of the model being flawed or too simple for the task.

The designs of the two types of contact simulation ("pidgin" and "creole") are extremely simplistic. Agents with different initial languages are simply grouped together and interact with each other developing a new shared system. The concept of a superstrate, teaching its language to agents with different systems, who are communicating amongst themselves, was not implemented, and contact only occured between agents from two different systems and not more. These implementations were intitially left out to examine the behaviour of the model in a more basic contact situation without, unnecessarily, complicating it.

One feature in the model which seems to make the resulting vowel systems of the contact simulations have more vowels than we would expect in real contact languages, is that it is much easier for an agent to gain a vowel than to lose one that is already an intergral part of its space.

The clean up process employed in the model occurs on the two agents involved in the game with a probability of 0.1 each time. It removes any vowel from the agent's itinerary if it has been used five or more times and if it has a vowel success ratio below 0.7 (meaning that is has been unsuccessful in 30% of its games).

Therefore, if a vowel is successful over a very large number of games it will require a vast quantity of unsuccessful games for it to eventually be deleted. For example, if a vowel had been used in 500 games and had been successful in 450 of those (the success ratio would be 0.9) it would require a further 150 unsuccessful games before it could be deleted.

Adding a vowel, however, can happen much faster. It can, in fact, happen after only one imitation game. This is because of the process which adds a similar vowel to an agent's space if an unsuccessful game is achieved with a successful vowel. If an agent, assigned to the imitator role, chooses a vowel from its space which has a high success rate for the games and the game is a failure then this agent will create a new vowel in its space which, when synthesised, is close to the acoustic signal as is possible. This is done to avoid an agent moving a vowel in its space which may be an established vowel in the population. This feature of the model makes it very easy for an agent to adopt a new vowel which may turn out to be successful if it is shared across enough agents in the population.

It is, therefore, not surprising that agents in these contact simulations will converge on a system which seems large as they will most probably adopt any vowel which is, to some extent, well integrated among at least a few agents.

The lower average number of vowels per agent in the "creole" simulations may be explained as simply being the result of the new agents not having had enough time to aquire all the vowels in the time they have been given.

de Boer (2001) discusses that the vowel systems in populations with blank agents constantly added, converge on configurations with fewer numbers of vowel clusters than in static populations.

The model, therefore, is not suitable to replicate the findings from real pidgin and creole languge phonology. But if the model can be used to successfully illustrate the emergence of vowel systems in populations of agents, why can it not model the emergence of contact vowel systems?

Perhaps contact languages are too complex to model at all.

6.1.2 Contact languages are too complex

Another justification for the shortcomings of the simulations, other than declaring the model unsuitable, is that contact languages, like pidgins and creoles, are simply too complex to replicate in a computer model.

Why do real contact languages seem to adopt such a simple set of vowels when, generally, they are in contact with a great deal more (namely the sum of all the vowels in each language involved).

The answer may lie in the fact that the domain in which the pidgin is used is very limited. In scenarios such as trade (where we often find pidgins being used), there is no need for the tradesmen speaking the pidgin to be able to discuss a huge variety of matters. Hence, they will only need a limited number of words in their vocabulary. It, therefore, seems a reasonable hypothesis that in order to be economical and not have more phonemic distinctions than they require to make the words in their vocabulary distinguishable from each other, they would not use a wide variety of sounds.

Nowak & Krakauer (1999) suggest that the "probability of misunderstanding a signal [...] limits the number of objects than can be described by a protolanguage" in their model of language emergence in a group of agents. Zuidema & Westermann (2001) discuss the "optimal lexicon" which yields the highest communicative success and, therefore has specificity (every meaning has exactly one form to express it and every form has exactly one interpretation), distinctiveness (used forms are maximally dissimilar to each other as to be easily distinguished) and sharedness (all agents use the same forms for the same meanings).

These two notions can both be linked to the idea of a communication system, with very few meanings to convey, only requiring a limited number of symbols to convey them. If a protolanguage was initially likely to only be able to convey a limited set of meanings due to the fact that a limited number of signals were available then the opposite must have been true. Having very few meanings to convey only requires a limited number of signals.

The combination of signals can permit a limited set of symbols to convey more complex meanings (Kirby 2000, Nowak & Krakauer 1999), but there is no need for a vast inventory of signals to be used to convey only a very small number of meanings. This would not adhere to the notion of specificity that Zuidema & Westermann (2001) claim is required for a system with high communicative success.

Extending the idea to phonology it, therefore, seems likely that a communication system with a a limited number of meanings, and hence words, will only require a limited number of phonemes to represent the words

A limited set of meanings to convey could therefore explain the limited phonology we find in pidgin languages.

The model used for this study cannot take this idea of limited domain into consideration.

In this model there is almost no semantic represenatation whatsoever. The only aspect that could be considered semantic to some degree is the agents' ability to decide whether a vowel uttered back to them is “right” or “wrong”.

Another point of interest in the study of contact linguistics is the social factors affecting them. This is often the main focus when examining contact languages (see Winford 2003, Thomason 2001, Sebba 1997, McMahon 1994, Romaine 1988, for example).

Implementing sociological factors in a computer model would, however, be extremely difficult and most probably unrealistic.

6.2 Other linguistic phenomena concerning the interaction of phonological systems which could be investigated with de Boer's model

6.2.1 Accent alteration

It is not uncommon for a speaker to want to change the accent of his/her language. This could be, for example, a temporary change when playing a part in an acting role or could be an attempt to fit in with a new community after moving to a new gegraphical location.

The difficulty in acquiring a new accent can vary greatly. If the new accent the speaker is attempting to acquire has all the same phonemes as his/her previous but simply uses them in different contexts, then it would be, in principle, relatively straight forward for the speaker to acquire the new accent. However, to go from having a single phoneme for a certain context in the original accent to having to split it into two seperate phonemes for different positions in the accent that is being acquired is much more difficult.

"Where the difference between the original accent and the target accent is a systemic one, the adaptation is easy if the original accent overdifferentiates with respect to the target accent, but fraught with difficulty if it underdifferentiates."(Wells 1982, p.113).

Wells (1982) discusses a difference between Standard Scottish English (SSE) and GenAm (General American) to illustrate his point. GenAm has two phonemes, /u/ and /ʊ/, where SSE has only one /u/. GenAm, therefore, overdifferentiates with with respect to SSE meaning that speakers of GenAm should, in theory, not have any problems in producing a SSE accent, since all they have to do is simply substitute the phoneme /ʊ/ they would use in their original GenAm accent with the vowel /u/ which they already have. A SSE speaker trying to achieve a GenAm accent, however, will encounter more difficulties as he or she will have to split the /u/ phoneme in two, learning that in some words (like good and put) in must altered to /ʊ/, whereas in other it should be kept as a realisation similar to /u/.

I believe that attempting to simulate the phenomenon of accent alteration using de Boer's model would be an interesting possibility. The coming together of vowel systems in this project was interpreted as contact of two different languages but the vowel systems could just as easily be interpreted as the vowel itineraries of different accents rather than languages.

It would be a rather difficult scenario to model successfully, but I believe it may provide some interesting results.

Wells discusses the fact that acquiring a new accent is generally a matter of adding new phonological rules to a phonological repertoire:

"...it does not merely exploit what are already the latent capacities of our built-in variability, acquiring a new accent can be seen as consisting mainly in adding new, late rules to our existing phonological competence." (Wells 1982, p.111)

The fact that the results of this study seem to show that agents in the model can aquire new vowels a great deal faster than they lose them seems to, somewhat agree with this detail.

It does seem plausible to suggest that an adult speaker learning a new accent will be able to learn a new phoneme present in the accent being learned (even though they may find it difficult and the realisation of the phoneme may not be quite the same as that of a speaker who has used the accent from an early age) much faster than the time it would take to lose the phonemes from their original accent not present in the new accent.

6.2.2 *Second Language Acquisition*

It is possible that where de Boer's model fails in simulating the vowel systems we find in real life contact languages it could succeed in explaining some of the aspects associated with second language (L2) learning. Suter (Suter 1976, Purcell & Suter 1980) has shown that the accent of second language English speakers varies according to their native language with, in his experiment for example, Persian and Arabic people speaking English as an L2 using much more "English-like" accents than Japanese and Thai people speaking English as an L2. Suter (1976) suggests that this could be due to phonological interference assuming that “structures of the Arabic and Persian languages facilitate the pronunciation of English to a much greater extent than do the phonological structures of Japanese and Thai” (Suter 1976, p.248).

Piske & Flege (2001) appear to agree with these claims: “It seems reasonable to assume that the accuracy with which non native speakers pronounce an L2 is, at least to some extent, dependent on their L1”(Piske & Flege 2001, p.193).

The study from which this quote is taken also illustrates the importance of the amount that the L1 and L2 are used (Flege & Nozawa 1997, also), and the age that L2 is first acquired, in predicting how native-like the accent for the L2 will be.

An experiment investigating the English accents of Italian-English bilinguals living in Canada (Piske & Flege 2001) emphasises the importance of the age of acquisition, with speakers who acquire English at a younger age having a better English accent, aswell as the importance of the amount of L2 spoken, with speakers who speak English most of the time, rarely speaking Italian, having the best accent .

“native speakers of Italian who continue to speak their L1 frequently have significantly stringer foreign accents in English than do individuals who speak their L1 infrequently.” (Piske & Flege 2001, p.209).

It would be interesting to see how well the model put forward by de Boer could measure the effects that the structure of the first language and the amount to which it was used have on the accent of a second language.

We could imagine a scenario in which an agent who has “grown up” in a simulation with a certain vowel system is then included in another group of agents who have been involved in a different simulation and hence who have a different vowel systems, while at the same time, still interacting with the agents he was initially with. The agent’s personal vowel space could be examined, checking to see whether the initial structure of his vowel space or the percentage of time spent with the new population compared to the old one has any interesting effects on it.

6.3 Conclusions

To conclude we must note that the results of the simulations performed, examining situations of basic language contact, do not match the patterns which are discussed in the contact linguistics literature. Agents' ability to gain vowels easily makes the possibility of a reduced system, as we find in real pidgins, resulting from a situation of contact between populations with different vowel systems, highly unlikely.

The fact that the model fails should not, however, lead us to the conclusion that a genetically determined language device is at work when humans create pidgin and creole languages. To model the phenomenon of language contact, a more complicated computer architecture needs to be used, which can implement more of the factors important in the language's birth. These could include the implementation of some sort of semantic context or a more complicated sociological arrangement.

Though this would be a difficult task, as it would be quite easy to over complicate the model, almost manufacturing the results that emerged, it may give us some insights to the question of why contact languages are the way there are.

References

Appel, R. & Muysken, P. (1987), *Language Contact and Bilingualism*, Edward Arnold.

Bender, L. M. (1987), Some possible african creoles: a pilot study, *in* G. Gilbert, ed., 'Pidgin and Creole Languages: Essays in Memory of John Reinecke', University of Hawaii Press.

Bickerton, D. (1981), *Roots of Language*, Karoma.

Bickerton, D. (2001), How to acquire language without positive evidence: What acquisitionists can learn from creoles, *in* M. DeGraff, ed., 'Language Creation and Language Change: Creolization, Diachrony, and Development', MIT Press, pp. 49–74.

Chomsky, N. & Halle, M. (1968), *The Sound Pattern of English*, MIT Press.

de Boer, B. (2000), 'Self organization in vowel systems', *Journal of Phonetics* **28**(4), 441–465.

de Boer, B. (2001), *The Origins of Vowel Systems*, Oxford University Press.

Flege, J. E., F. E. M. & Nozawa, T. (1997), 'Amount of native language(l1) use affects the pronunciation of an l2', *Fournal of Phonetics* **25**, 169–186.

Gonzales, A. (1922), *The Black Border: Gullah Stories of the Carolina Coast*, State Company.

Hall, R. A. (1966), *Pidgin and Creole Languages*, Cornell University Press, Ithaca.

Holm, J. (1988), *Pidgins and Creoles*, Vol. 1: Theory and Structure, Cambridge University Press.

Holm, J. (2000), *An Introduction to Pidgins and Creoles*, Cambridge University Press.

Kirby, S. (2000), Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners, *in* H. J. Knight, C. & M. Studdert-Kennedy, eds, 'The Evolutionary Emergence of Language: Social function and the origins of linguistic form', Cambridge University Press.

Ladefoged, P. & Maddieson, I. (1996), *The Sounds of the World's Languages*, Blackwell.

Liljencrants, J. & Lindblom, B. (1972), 'Numerical simulation of vowel quality systems: The role of perceptual contrast', *Language* **48**(4), 839–862.

Lindblom, B. (1986), Phonetic universals in vowel systems, *in* J. J. Ohala & J. J. Jaeger, eds, 'Experimental Phonology', Academic Press, Inc., pp. 13–42.

Lindblom, B. & Lubker, J. (1985), The speech homunculus and a problem of phonetic linguistics, *in* V. A. Fromkin, ed., 'Phonetic Linguistics: Essays in Honor of Peter Ladefoged', Academic Press, Orlando, pp. 169–192.

Lindblom, B., M. P. & Studdert-Kennedy, M. (1984), Self-organizing processes and the explanation of language universals, *in* C. B. Butterworth, B. & O. Dahl, eds, 'Explanations for Language Universals', Berlin: Walter de Gruyter & Co., pp. 181–203.

McMahon, A. M. S. (1994), *Understanding Language Change*, Cambridge University Press.

Mühlhaüsler, P. (1986), *Pidgin and Creole Linguistics*, Basil Blackwell.

Nowak, M. A. & Krakauer, D. C. (1999), The evolution of language, *in* 'Proceedings of the National Academy of Science: USA', Vol. 96, pp. 8028–8033.

Oudeyer, P.-Y. (2006), *Self-Organisation in the Evolution of Speech*, Oxford University Press.

Piske, T., M. I. R. A. & Flege, J. E. (2001), 'Factors affecting degree of foreign accent in an l2: a review', *Journal of Phonetics* **29**, 191–215.

Purcell, E. T. & Suter, R. W. (1980), 'Predictors of pronunciation accuracy: a reexamination', *Language Learning* **30**(2), 271–287.

Romaine, S. (1988), *Pidgin and Creole Languages*, Longman.

Schwartz, J.-L., B. L.-J. V. N. & Abry, C. (1997), 'Major trends in vowel system inventories', *Journal of Phonetics* **25**, 233–253.

Sebba, M. (1997), *Contact Languages: Pidgins and Creoles*, MacMillan Press.

Steels, L. (1995), 'A self-organizing spatial vocabulary', *Artificial Life* **2**, 319–332.

Steels, L. (1996), Emergent adaptive lexicons, *in* M. M. J.-M. J.-A. P. J. Maes, P. & S. W. Wilson, eds, 'Proceedings of the International Conference on Multi-agent Systems', MIT Press, pp. 562–567.

Suter, R. W. (1976), 'Predictors of pronunciation accuracy in second language learning', *Language Learning* **26**(2), 233–253.

Thomason, S. G. (2001), *Language Contact*, Edinburgh University Press.

Vallée, N. (1994), Systèmes vocaliques: De la typologie aux prédictions, PhD thesis, Institut de la Communication Parlée (Grenoble-URA C.N.R.S. no 268).

Wells, J. C. (1982), *Accents of English 1: An Introduction*, Cambridge University Press.

Winford, D. (2003), *An Introduction to Contact Linguistics*, Blackwell.

Zentella, A. C. (1997), *Growing Up Bilingual: Puerto Rican Children in New York*, Blackwell.

Zuidema, W. H. (2005), The Major Transitions in the Evolution of Language, PhD thesis, University of Edinburgh.

Zuidema, W. & Westermann, G. (2001), On the relevance of language evolution models for cognitive science, *in* 'Presented at DECO-2001, Edinburgh'.