Radboud University of Nijmegen

M.Sc. Thesis

How to-mah-to became to-may-to:

Modelling and simulating linguistic dynamics

Author:

Bastiaan du Pau (s0368628)

Supervisors:

Prof. dr. Ton Dijkstra

Dr. Pim Haselager

Dr. Walter van Heuven

Reading committee:

Prof. dr. Ton Dijkstra

Dr. Pim Haselager

...

July 18, 2010

Contents

List of Figures

List of Tables

Acknowledgements

Pim Haselager was the first to see potential in this project when I wrote a proposal for it during one of his courses, and it motivates me to feel he still sees it. I deeply appreciate the mystical and musical topics that passed in conversations with Ton Dijkstra and the effort I know he puts in supporting my work. If it wasn't for my supervisors Ton and Pim telling me to become more concrete and start finishing the project, I would still be in the exploratory phase while you read. Thanks for teaching me that every product is a work in progress. Walter van Heuven was a great host at the University of Nottingham, where I implemented the simulation tool for this thesis. Our weekly meetings and his eye for results helped shape the project in the early stages. I owe gratitude to Paul Kamsteeg for finding the time to review the initial design of the simulation tool. His remarks took the design to a next level. Big up for Job Schepens, with whom I shared a house and office in Nottingham. Besides learning life lessons together, I value the abstract conversations we had (and still have) from time to time. Rik van den Brule and I fought side by side in the battle with LaTeX, which ended in a close victory. Since I prefer to avoid brain storms, I thank Peter Indefrey, Tom Claassen, Pieter Muysken, Michael Dunn, Fiona Jordan, Jules Ellis, Peter du Pau, Elmer Doggen, Joost van Doremalen, and Michael Franck for the delightful brain breezes we had.

Chapter 1

Introduction

Language is dynamic. In an analysis that made the national newspapers, Harrington (2000, 2005) revealed that HRH Queen Elizabeth II's vowel pronunciation shifted from a Received Pronunciation (RP) towards a 1980s standard southern-British accent (SSB). Although RP, ironically also known as Queen's English, is spoken primarily in the upper class and thereby grants its speaker a certain prestige, SSB is spoken typically by lower class youngsters. Acoustic comparison of 11 vowel sounds in the Queen's Christmas messages from the 1950s with those from the 1980s showed that there was significant change in 10 of these 11 vowel sounds. Even more interestingly, the Queen's pronunciation of these vowel sounds in the 1980s held somewhere in the middle between her old pronunciation and SSB, indicating that the Queen did not preserve her dialect but shifted it significantly in the direction of SSB.

This sort of phenomenon is not restricted to royalty. Imagine yourself conducting smalltalk with a group of people speaking a different dialect. How do you converse? Do you start sounding more like they do or do you remain persistent in speaking your own dialect, and if you do change your pronunciation, is this temporary or does it have a long-term effect? One other entertaining example of linguistic dynamics can be appreciated in a duet written by George and Ira Gershwin called "Let's call the whole thing off". In a famous recording of this song, Louis Armstrong and Ella Fitzgerald contemplate if they should stop seeing one another because of the differences in their dialects. Part of the lyrics famously points out the different pronunciations of tomato: to-may-to vs. to-mah-to. In the end, they seem to favour love over language, though. In this thesis you will read more about the mechanisms behind such vowel changes and differences as we model and simulate linguistic dynamics.

The goals of this study were threefold. The first goal was to develop a model of linguistic dynamics based on literature study. This resulted in a model that is both compact and extensive.

Three hypotheses, pertaining to the formation of a shared language, linguistic change, and linguistic divergence that emerged from it were:

Formation hypothesis reinforcement and memory decay drive lexical formation,

Change hypothesis alignment drives linguistic change, and

Divergence hypothesis geographical constraints drive linguistic divergence.

A brief explanation of each of these hypotheses is given in the following paragraphs.

Lexical formation is the state change of a population from being without language to having a complete, shared language. In contrast with current belief, the formation hypothesis did not contain synonym punishment as a necessary factor in lexical formation. Instead, punishment of disuse was expected to ensure unambiguity in languages, which itself arise from the reinforcement of associations after successful communication.

Alignment is a mechanism people use either unintentionally, to be socially accepted or to facilitate communication. It involves adjusting your speech in the direction of the person you interact with (your interlocutor). This mechanism has been shown to exist on several linguistic levels, ranging from phonology to syntax and situation model. If the internal lexicon is affected by the speech adjustment, subsequent utterances will also be influenced and linguistic change is a fact.

Geographical constraints drive linguistic divergence because they contribute to the fragmentation of a population into smaller communities that interact more with each other than they do with other communities. An example of this the Great Channel, which separates Great Britain from Northern France. Besides oceans, other examples of constraining geographical features are deserts, mountains, rivers and cliffs. This, assuming that the change hypothesis holds, leads to the communal languages changing in independent directions and in doing so diverging from each other. A more detailed description of the model can be read in Chapter 4.

The second goal was to design and implement a tool with which linguistic dynamics can be simulated. One additional aim was that the tool can be used by other scientists for further investigation of this model and dynamics of language in general. The outcome of this goal was the Dialect Emergence Virtual Lab (DEViL). This simulation tool uses a new type of language game as local interactions from which the global properties emerge, called the aligning game. This game is based on the guessing game as employed in the Talking Heads experiment (Steels, 1999). The two important changes with respect to the guessing game are the presence of alignment and the absence of synonym punishment. In the DEViL, software agents are able to walk through a map of locations, and to play aligning games with agents that are present in the same location. The agents communicate using a lexicon that links vowels directly to local objects; an agent thus speaks in vowels, such as /a/, /i/, or /u/. Details of the tool can be read in Chapter 5.

The third goal was to use this tool to provide simulation evidence for the model of linguistic dynamics. This was a challenging but important goal, because modelling the evolution of ongoing linguistic change is one of the key open challenges in language evolution modelling put forward by de Boer and Zuidema (2010). To achieve this goal, one simulation was set up for each of the three hypotheses stated above. The idea that learning and memory decay underlie the formation of a shared language was backed up by a simulation in which agents attempted to develop a shared lexicon under different degrees of reinforcement and decay. Only when both reinforcement and decay were present, the agents had a chance of succeeding in this attempt. In a second simulation, a community of agents interacted for a prolonged period of time under varying amounts of alignment. The communal language was shown to change with a speed roughly proportional to the degree that agents aligned their speech to the agents they interacted with. In the last simulation, the effects of geographical constraints were modelled by varying the travel rate in a simulation in which two communities with the same language were put in two neighbouring locations. Depending on the amount of travelling, the communal languages either diverged to become unintelligible languages or intelligible dialects, or remained identical throughout the simulation. This is in accordance with both the model's predictions and reality.

It further appeared that various simulation parameters affected the chance of lexical formation succeeding. For example, larger communities create a bigger number of synonyms in the early phase of the formation of a shared language, which impedes formative success. Yet, in subsequent phases, if one of the synonyms starts to become the leading candidate for being used in the shared language, it is more quickly reinforced in larger communities than in smaller ones, which assists formative success. Acoustic noise was also shown to have multiple effects on the dynamics of language. One of these effects was that it forced communities to use the vowel space optimally. Vowel systems in the real world often show optimal use of vowel space as well, indicating just one of the interesting positions noise takes in linguistic dynamics.

The analysis of the simulation results lead to several innovations regarding the initial model. Noise was embedded in it as it proved to have several effects. Four phases in the formation of a shared language were identified, namely the creation, propagation, competition, and maintenance phase. Also, distinctions between states that two neighbouring languages can be in were made. Finally, a more detailed model of linguistic change and divergence allowed us to predict when and which state transitions would occur based on rates of travel.

The thesis is arranged as follows. In Chapter 2, vital concepts such as language as a complex adaptive system, the vowel space, the language game and alignment are explained. A summary of seminal and recent AI research on the topic of language evolution is given in Chapter 3. In Chapter 4, one can read about the linguistic dynamics model that is introduced in this thesis.

In the basic simulation chapters, Chapter 6, 7 and 8, three hypotheses that come forth from the model are tested with the DEViL. Specifically, the simulation in Chapter 6 tests if reinforcement of communicative success and punishment of disuse drive lexical formation, the simulation in Chapter 7 tests if vowel change is driven by alignment, and the simulation in Chapter 8 tests if geographical constraints drive vowel divergence.

A deeper analysis of these simulations' results is given in Chapter 9. In this chapter, the two main topics under revision are the parameters affecting the successful formation of a shared language and the effects of acoustic noise on linguistic dynamics. Chapter 10 presents a general discussion of the results' implications and interesting options for future research. Concluding remarks can be found in Chapter 11.

Chapter 2

From language to vowel change

This chapter provides the reader with theoretical backgrounds involving language and evolutionary linguistics in general, and the principles and causes of vowel change. First, a framework for language is presented centering around the idea that the utterance is the basis of language in use and the view of language as a complex adaptive system. Vowel change is explained and examplified thereafter. Last, three potential causes of vowel change are given: vowel space asymmetry, alignment and misperception of co-articulation.

2.1 Language

Language has claimed such omnipresence, that it is hard to imagine life without it. Words are always buzzing in our heads; they fill our rooms either through the bellowing of our contemporaries or the speakers of our full HD televisions. Our thoughts would be as concrete and direct as those of a baby if it was not for language. In this section, I focus on explaining the most relevant concepts pertaining to language used throughout the thesis. These include the semiotic square and the vowel space. But doing so requires a definition of what language actually is.

Giving a definition of language is not straightforward. One definition that is very elegant and also useful in the context of this thesis is given in Labov (1994) and goes as follows: language can be defined as a system of associations between signals and meanings used as communication by a community. There are other aspects of language such as syntax, the study of how words are combined to construct sentences, but these currently have less relevancy for this thesis. In semiotics, de Saussure and Peirce proposed the same association between the signifier (signals, for example utterances or gestures) and signified (meaning or mental concept). The linking of signals to meanings does not happen in the air: it happens in people. That is why Peirce expanded the relation to

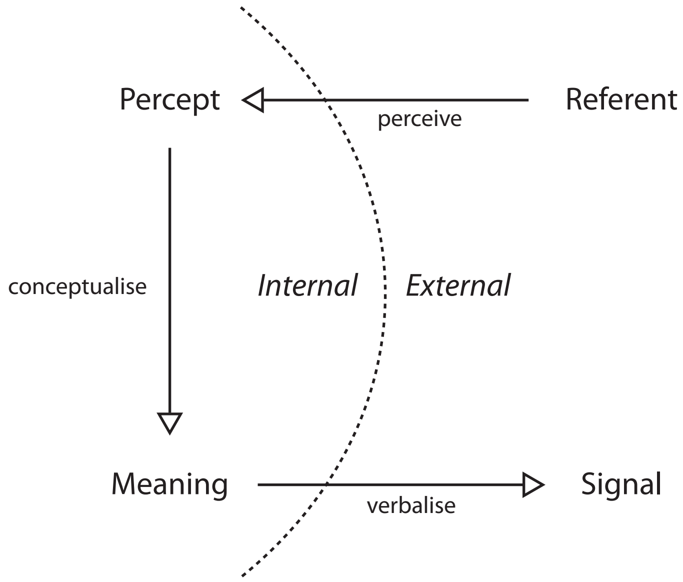

Figure 2.1: The semiotic square: a visual model of language in use.

a triadic one, including also the interpretant (Liszka, 1996). The interpretant is included in the above definition in the form of the community using the system of associations. A pleasant way of visualising and modelling language is via the semiotic square.

2.1.1 Semiotic square

The semiotic square, depicted in Figure 2.1, is a visual model of language in use. The bottom part shows the association between signal and meaning that is the basis of a language. The signal is outputted into the world through means as communication, articulation. Meaning is linked to objects in the world through observation, experience and conceptualisation. This path is portrayed by the upper half of the image. In humans or agents, a set of relations between percept and meaning is typically called an ontology, while one between meaning and signal is called a lexicon.

The left part is located within the user of a language. The user cuts up the world into categories, labels them and lets them speak. As such, an infinite, continuous world is transformed into discrete, hierarchical elements by means of language. This categorisation is paramount in enabling us to think, reason and communicate about the world. A side effect of this is that language, in doing so, also shapes the way we perceive the world. For example, members of the Pirahă tribe in Brazil tribe, which has a language with words for one, two and many instead of a full numbering system, have been shown to have difficulties in recalling numbers larger than three (Borensztajn, 2006). This effect is captured by the much debated Sapir-Whorff hypothesis.

Anything can act as a signal in a language, as long as it carries information perceivable to the intended audience. In human communication, speech is often used as carrier of linguistic signals.

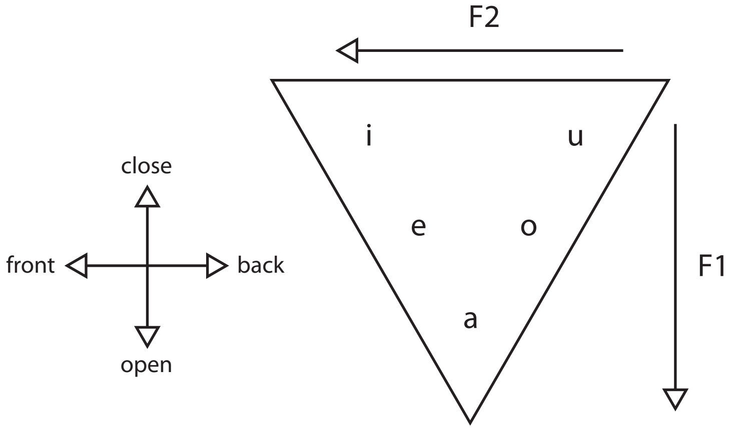

Figure 2.2: The vowel space. Note that close/open is also called high/low and that front is sometimes called bright.

Other media can and have been used though, such as sign languages, writing, Braille and Morse code. These all arose in cases where speech is impossible or insufficient, namely when one can not hear or see, or when one is too far away to hear. This indicates that in humans, speech is indeed an the primary candidate to carry signals. In the next section we treat this phenomenon called speech and, more specifically, the vowel space.

2.1.2 Vowel space

Within speech, linguists make a clear distinction between two groups of phonemes: vowels and consonants. The difference between the two is that the vocal tract is open for vowels and partially or completely closed for consonants. How vowels and consonants combine to form fluent speech is not so straightforward. Two rough ways of thinking about this are the following. One can view speech as the rapid succession of vowels and consonants, thus forming words and sentences. Alternatively, speech can be seen as consisting of a continuous understream of modal voice, interspaced by different kind of interruptions of the airflow. Both views are abstractions of what happens in reality and thereby serve a purpose in making it easier to talk about speech in sensible terms. They distract from the fact that the difference between vowel and consonant is not always clear; also time boundaries between them are hard to exactly pinpoint.

Vowels can be characterised and thus distinguished by their productive features, including but not being restricted to height, backness and nasalisation. Figure 2.2 depicts the vowel space that is spanned by the two features height and backness. Vowel height is called after tongue position, high vowels having the tongue positioned high in the mouth, low vowels having it positioned low. Vowel openness is essentially the same as vowel height: it is called after jaw condition, the opening the jaw leading to a similar acoustic result as lowering the tongue. As can be seen in the figure, /i/ and /u/ are typical high or close vowels: they are pronounced with the tongue relatively high or the jaw relatively closed. /a/ is a typical low or open vowel: it is pronounced with the tongue low or the jaw open.

Vowel backness is also called after tongue position, back vowels having the tongue positioned to the back fo the mouth (but not so far that it touches the soft palate and becomes a velar consonant), front vowels having it positioned to the front (but not so far that it touches the teeth or lips and becomes a dental or lingolabial consonant). Vowel brightness is essentially the same as vowel backness: it is called after the sound they make. As can be seen in the figure, /i/ is a typical front vowel: it is produced with the tongue to the front. /u/ is a typical back vowel: it is produced with the tongue to the back.

You can characterise and distinguish vowels by their perceptive features as well. Most often used as features are so-called formants, but others can be thought of, like amplitude. In speech, formants are resonances in the vocal tract that lead to peaks in the voice's sound spectrum. Besides the fundamental frequency (f_0, which is low for big, grown men and high for little girls), the formant with the lowest frequency is called f_1 and each formant with a higher frequency gets a higher number. As can be seen in Figure 2.2, the first two formants can be used to distinguish vowels. Roughly, lowering the tongue increases the frequency of the first formant f_1 and bringing the tongue to the front increases the frequency of the second formant f_2. The same vowel space can thus be said to be spanned by either productive or perceptive features; is it not amazing how the speech and hearing organs have coadapted? It should be noted that a change in tongue position does not coincide with the same change in frequency; instead this relation is disproportional and rich of plateaux where the frequency is relatively stable given small changes in tongue position (Lieberman, 2006).

/i/, /a/ and /u/ are very common vowels: they appear in 87%, 87% and 82% of the languages in the UPSID_{451}, the UCLA Phonological Segment Inventory Database (Maddieson, 1984). This database contains phoneme inventories for 451 languages. In comparison, the vowel /y/, which is a sound holding in the middle of /i/ and /u/, appears in only 5% of the languages in the UPSID_{451}. In fact, Crothers (1978) made a typology of the world's vowel systems that grants /i/, /a/ and /u/ even bigger importance. Based on observing a database that is a predecessor of UPSID, Crothers created the formal vowel system hierarchy depicted in Figure 2.3. According to this hierarchy, any language with three vowels has /i/, /a/ and /u/. Each language with a fourth vowel either has /ɨ/ or /ɛ/. which the relative positions in the vowel space. Further down the line, /ɔ/, /e/ and /o/ are typically added. Note that Crothers intended to classify languages based on the relative position

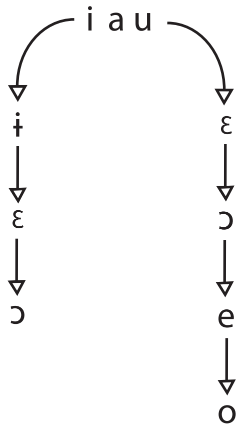

Figure 2.3: Part of Crothers' typology of vowel systems. /i a u/ appear in all vowel systems with three vowels. /ɨ ɛ ɪ e o/ are added in bigger vowel systems.

of its vowels, not necessarily the absolute position. This means that any system with three vowels forming an upside down triangle is classified as a system with /i/, /a/, and /u/, even though it may not necessarily contain these vowels as defined by for instance the IPA.

2.1.3 Complex adaptive system

The view of language as a complex adaptive system (CAS) is adopted more and more (Steels, 1999; The Five Graces Group, 2010). It contrasts, as do more models and theories we come across in the continuation of this chapter, with the traditional static view of language of the generativist approach. In doing so, it bridges the gap with this section that is concerned with language and its properties, and the next section that is more concerned with the dynamics of language. These two are more intimately related than previously thought.

The four key features of the CAS, as put forward by The Five Graces Group (2010) are the following. First, the system is built up of individuals interacting with each other. Second, their interaction is based on their history of past interactions. Third, several factors affect the speakers behaviour, such as perceptual mechanics and social motivations. Fourth, experience, social interaction and cognitive processes together give rise to linguistic structures. Although these features are abstract and generic, they fit well with the view held by computational modelers, as we will see when we review the Talking Heads experiment (Steels, 1999). The simulations done in this thesis also have a much stronger affiliation with the view of language as a CAS, as opposed to the generativist view. The results that come from it even argue for a rephrasing of the last feature as "give rise to linguistic structures, change and divergence".

2.2 Vowel change

Language is dynamic. Languages are constantly changing, resulting in phenomena like chain shifts and dialect continua. Yet, this dynamic aspect is not at all apparent to the general public. Most changes (besides obvious generational trends of words for super, cool and darn) are either so small that they lie below the level of public awareness or so slow that they span centuries. People perceive a change as being constant because they themselves change along with it, just as a boat flows along with a river. This prevents people from always feeling uncertain of the state a language is in, but as a side effect keeps them from noticing the change.

On top of that, dictionaries and other efforts made to identify parts of a concrete language, like Latin or Modern English, make think otherwise by attempting to approach a static language. It is said that before this standardisation, one could travel on foot from the Netherlands to Italy without ever having to learn a new language. Stopping every few miles, one adjusts to slight changes towards a new dialect while talking to new people; thus accumulating gradual changes one ends up speaking the native tongue wherever one travels. Whether this is true or just fantasy is impossible to tell, the idea by itself is at least romantic and entertaining.

In this section, the dynamic nature is described and examplified. First, vowel change is placed in the general field of evolutionary linguistics. After this, examples of historical and current vowel changes are given. Then, the reader is given an account of vowel change principles, after which the section is concluded by the description of three mechanisms that are thought to be potential sources of vowel change.

2.2.1 Language evolution

Evolutionary linguistics, a hot topic, aims to identify when, where, and how a language emerges, changes and dies out (Ke and Holland, 2006; Gong, 2010). The first thing you find out when delving into language evolution is that it is a hard problem to solve because of the severe lack of data. Christiansen and Kirby (2003) even write it is the hardest problem in science! The historic origin of language cannot be observed because it is not in the present and does not repeat itself (Cangelosi and Parisi, 2002). Typically, the evidence from which to build theories on is scarce and buried deep under the ground, only to be excavated by archeologists. On top of that, the nature of the evidence, whether bones, manuscripts, broadcasts or experimental recordings often differ from each other, making comparison tricky. Over a century ago, this even made the Société de Linguistique de Paris famously place a ban on researching language evolution. Only since the 1960s has the field been reborn again. A paper by Steven Pinker and Paul Bloom is considered a catalyst for the field in

1990, effectively increasing the number of papers on language evolution tenfold (Christiansen and Kirby, 2003).

Evolutionary linguistics is not only a hard problem, it is also multifaceted. Many different subfields of research fall under its reign. Kirby (2002) made a distinction between three adaptive systems, namely phylogeny or biological evolution, glossogeny or language change and ontogeny or individual learning, which can be loosely related respectively to language origins, change and acquisition. Whereas phylogeny works on a time scale of millions of years, glossogeny and ontogeny work on centuries and even decades. Neither process should be underestimated, since they are all very complex. More about Kirby's work is discussed in the section about the Edinburgh experiments. Besides the type of system, research within the field of evolutionary linguistics varies along several other parameters. To name one, the level that language acts on can range from the individual, via the communal and the racial to the general. In Chapter 3, computational models of evolutionary linguistics are identified according to these distinctions.

Plenty of studies have been conducted on the topic of language evolution. Unfortunately, this thesis is not the right place to go through them all. Among many recent, interesting research are Senghas et al.'s (2010) investigations of Nicaraguan sign languages and the evolution of the larynx and (the absence of) air sacs (de Boer, 2010; Hombert, 2010). Another interesting field of research is that of Pidgin and Creole languages (Hall, 1966; Bickerton, 1975), of which we will encounter computational models in Chapter 3. We now continue discussing principle of vowel change.

2.2.2 Principles of vowel change

In a seminal body of work, Labov (1994) discovered that most of linguistic change is governed by general principles, taking the formalisation of this process of linguistic dynamics to a next level. Although the proposed principles are mostly based on vowel change data and are meant to explain vowel change, some of the more basic ideas behind them can be generalised to other linguistic domains that are subject of change, such as syntax and prosody. To illustrate this, let us look a little deeper into some of examples of vowel change and see how Labov would have described them. Three more or less well known phonetic changes in the history of the English language are described and analysed briefly: the Great Vowel Shift, Labov's New York City warehouse investigations and the Queen's shift of pronunciation. After that, we will get back at the principles and try to find out why they exist.

In the nineteenth century and the first half of the twentieth, a British r-less pronunciation was the norm in New York City. The r-less pronunciation contradicts with r-ful pronunciation in words like bar, which are pronounced baa and bar in r-less and r-ful pronunciation respectively. In the years following the second world war, r-ful reappeared as a prestige variable in New York City speech. In the classic study of Labov in 1962 and a restudy by Fowler in 1986, people of various ages and social backgrounds were recorded in three different warehouses, Saks (expensive), Macy's (medium) and S. Klein (cheap), answering "fourth floor" to a question asked by Labov. For each participant, it was noted if r was pronounced in the answer or not. It appeared that in 1962, r-ful was most prevailing in people visiting Saks, especially in young people.

The difference in pronunciation between generations interviewed simultaneously is an example of a change in apparent time. Although no direct information about a change at hand is available, one can say that people that were born later have a different pronunciation. With just the single measure, it is not deducible whether this is due to an actual sound change or simply because of age degradation. This is where Fowler's restudy comes in. She found that a quarter of a century later, r-ful had increased in popularity across all warehouses. This increase is known as a change in real time: when you interview the same person or type of person twice at different points in time, you can see if a change occurred in the meantime. In this case, r-ful was a real sound change, innovated by upper class youngsters.

The reappearance of r-ful is a typical example of change from above, which is introduced by the dominant social class (which in this case visited Saks). Changes from above first appear in formal speech and then either integrate with casual speech or form a separate speech form only used in formal contexts. Contradicting are changes from below, which can be introduced by any class and are thus not driven by social factors. These changes are, unlike changes from above, hardly ever noticed until they are complete. Linguistic change in general has proven in these and other studies by Labov to follow an S-shaped curve in time, changes starting out slow, reaching a maximum in midcourse and slowing down near completion.

The Great Vowel Shift took place in England from 1450 to 1750, changing the Middle English vowel pronunciation to a Modern English one. Whereas the Middle English pronunciation was similar to that seen on the continent, which makes sense historically, given both the shared ancestral backgrounds and the amount of invasions from the mainland, the Modern English variant three centuries later sounds much like, as its name suggests, modern English. An extended vowel chain shift consists of several successive minimal chain shifts. Each minimal chain shift takes two phonemes. The one vowel moves away from a spot in the vowel space that is occupied by the other vowel. This chain reaction phenomenon is typical of vowel shifts. In the case of the Great Vowel Shift, the two prominent extended chain shifts that occurred are shown in Figure 2.4. The one in the left should be

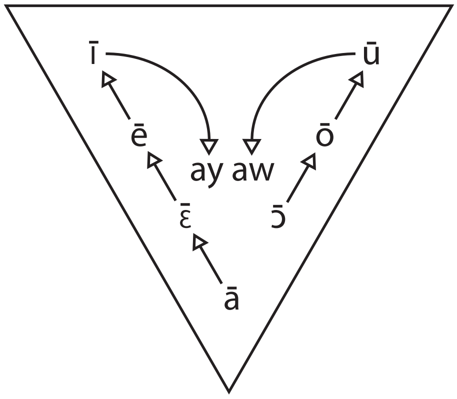

Figure 2.4: The Great Vowel Shift displayed in the vowel space. An extended chain shift moves upwards to /i/ and /u/, which themselves diphthongise to /ay/ and /aw/.

read as follows: the low, long vowel /ā/ rises to the position of /ɛ̄/. This vowel in turn changes to the position of /ē/, which in turn changes to the high, long vowel /ī/. It then becomes an upgliding diphthong /iy/, which nucleus falls from /i/ to /a/. A diphthong is a combination of two vowels within a syllable, consisting of one vowel that is most prominent, called the nucleus and one that it glides to or from, called the off-glide. So, time used to be pronounced like modern team, feet like fate, raid like red and can like Genghis Khan. The latter one is an exercise for the reader.

These chain shifts confirm two of the three general principles that Labov observed to govern chain shifts. These three are:

Principle I Long vowels rise

Principle II Short vowels fall

Principle IIa The nuclei of upgliding diphthongs fall

Principle III Back vowels move to the front

Note that long vowels are pronounced for a longer period of time than short vowels (think of four vs. for). All the changes except the ones moving down in Figure 2.4 are examples of Principle I, since /ā/ and /ɔ̄/ are long, low vowels and their destinations are long, high vowels. Common knowledge tells us that moving from a low position to a high one is rising. After diphthongisation, the two remaining changes are examples of Principle IIa; the nuclei of /iy/ and /iw/ are both high, while the nuclei of their destinations /ay/ and /aw/ are low, so they have fallen. These principles have proven to be quite robust, in that most shifts abide by them. Especially Principle I sees little exceptions. Besides vowel shifts, other phenomena of linguistic change are described in Labov (1994), such as splits and mergers. I strongly encourage the reader to read the book if the reader is in any way interested; it is a good read.

Coming back to the Queen's shift of dialect, described in the beginning of this thesis, we can now categorise this vowel shift on the individual level as being a change from below and in real time. It is from below because Queen's English is much more prestigious than the SSB accent, which the shift is moving towards. It is in real time, since multiple measurements at differents moments in time are taken of the same person, indicating that a real change has occurred in that person. This change is also an example of a gradual change, in contrast to a catastrophic one. Gradualism and catastrophism are both inspired by similar ideas in geology. The Earth can be shaped by gradual erosion or other forces, but also by catastrophic ones, such as floods. Similarly, language can also be affected by sudden events, such as invasions. Another important idea from geology that has taken the jump to linguistics is the uniformitarian principle. First proposed by Scottisch geologist James Hutton in 1785, it dictates that the forces that operated in the past are the same that operate today. Translated to linguistics, one arrives at the notion that the processes driving linguistic change in the present have always driven linguistic change. Now we have laid out the landscape of possible linguistic changes, at this moment we ask the question what actually drives it. In the following sections, I discuss three potential sources of change: vowel space asymmetry, alignment and co-evolution.

2.2.3 Potential sources of vowel change

The first source is vowel space asymmetry. Martinet (1955) came with an early account of sound change causality. In this account, two opposing tendencies are believed to be the driving force of sound change. The first is that speakers tend to preserve symmetry in a system: symmetric vowel systems are preferred over asymmetric ones. A symmetric vowel system typically has as many front vowels as it has back vowels. The second is the asymmetry of the articulatory space of the supraglottal tract: there is more space available for front vowels than for back vowels. Four degrees of height can be accommodated easily in the front, while this number leads to overcrowding in the back. These two tendencies clash. When a system has four front vowels, four back vowels are preferred for symmetry reasons, but this does not fit. The result is that the back vowels start shifting, causing vowel change.

Croft (2000) notes that teleological mechanisms such as this one are often not plausible. A teleological mechanism is one where sound change occurs because it is the explicit goal of the language users. Most of the times, it is unrealistic to credit people such goals. In case of the vowel space asymmetry mechanism, the speaker is said to have the explicit goal of preserving symmetry. Croft argues that speakers are not at all directly concerned with this symmetry.

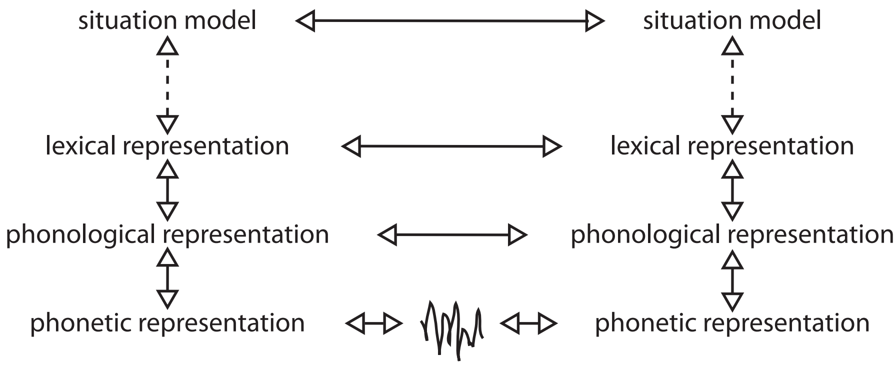

Figure 2.5: The Interactive Alignment Model. The right and left column represent two interlocutors. Horizontal arrows indicate channels that permit aligning. The bottom arrows that link the phonetic representations stand for verbal communication.

The second source is alignment, also called accommodation or convergence. As Pickering and Garrod (2004) point out, traditional psycholinguistic theory is based on monologue, while the most natural form of language use is dialogue. We have seen this discrepancy before when we discussed the view of language as a CAS. The main reasons for the neglect of dialogue are twofold. First, studying comprehension as opposed to dialogue is thought to be much more straightforward and experimentally controllable. Second, psycholinguistics is based on theories of isolated sentences without any context, developed by Chomskyan generative linguistics, which clashes with the interactiveness and richness of context found in dialogue. To move forward in understanding language use in general, Pickering and Garrod (2004) propose a model of dialogue processing called the interactive alignment model (IAM).

The IAM is an adjustment of the traditional view of dialogue in which multiple links are added between interlocutors beyond the usual link between the interlocutors' phonetic representations via the physical utterance. These links, which can be seen as dotted arrows in Figure 2.5, represent so called channels whereby priming occurs. The rationale behind these additional links comes from a series of recent experiments on dialogue, in which participants (often without themselves noticing) show alignment to their interlocutors on several levels of linguistic representation, ranging from the phonological representation to the situation model.

The latter is motivated by an experiment in which two participants independently look at a maze and interact about it to solve problems. Over time, the participants typically evolve into using the same representation of a situation model. For example, they both use coordinates or relative directions. Having the same representation is not a requirement for solving the problem, but it is what de Boer (2001) would call an attractor, because when you are aligned on the situation model level (or any level at that) you only have to use one model and have no need for conversion. This is computationally less heavy and is thus a more attractive state for the sytem of interlocutors and the problem at hand to be in.

Very similar to alignment is the notion of convergence in Howard Giles' communication accommodation theory (CAT) (Giles et al., 1991; Giles and Ogay, 2006). CAT was set up in the 1970s and provides a framework that predicts and explains the communicative adjustments people make with the purpose of creating, maintaining or decreasing social distance. Convergence is defined as a "strategy whereby individuals adapt their communicative behaviours in terms of a wide range of linguistic (e.g., speech rate, accents), paralinguistic (e.g., pauses, utterance length), and nonverbal features (e.g., smiling, gazing) in such a way as to become more similar to their interlocutor's behavior" (Giles and Ogay, 2006). Linking well with Labov's changes from above and below is the fact that convergence can be either upward or downward in societal valence. In an upward convergence, the individual adopts a more prestigious pronunciation. Note that frequent use of this strategy potentially leads to a change from above. Other strategies are possible, such as downward divergence, which occurs for example when a teenager puts emphasis on its street lingo while discussing with its parents.

Note that neither theory is teleologic: alignment is thought to happen either unconsciously and automatically or with purposes like enhancing communicative success (Lewis (1969) sees dialogue as a game of cooperation, where both participants win if both understand the dialogue and neither wins if not), seeking approval or altering social distance, not to change the linguistic system directly. This makes them more realistic than the vowel system asymmetry theory. Note also that neither theory mentions alignment to be the cause of linguistic change. The purposes of the theories are not change related, as comes forward from their definitions. Yet, alignment is a mechanism that potentially induces vowel change (and linguistic change in general), because the result of alingment is an alteration of the internal language. This is very clear in the case of vowels, when you align by shifting your vowels to that of your interlocutor, the vowels in your lexicon have changed position in the vowel space. For alignment to really drive change, the result of the alignment should be of a long lasting nature.

It is not known if alignment affects the internal language in a permanent way. Considering this, two scenario's can be thought of. On the one hand it could be so that contact with individuals with a different dialect temporarily changes your pronunciation, after which it returns to your habitual pronunciation. On the other hand it could be so that only the first happens and that your pronunciation remains changed. The latter is what Bybee (2006) and The Five Graces Group (2010) propose: according to them, every instance of language use changes the internal organisation. Of course, a scenario that holds in the middle of these two is also highly plausible; in such a scenario your habitual pronunciation moves a certain amount towards the altered pronunciation used during interaction with a different dialect.

The third source is misperception of co-articulation. This model of vowel change is slightly complex but very intriguing, and has been proposed by Ohala (1993). It was inspired by the attempt of Hiroya Fujisaki to re-unify the phonetic sciences and brings together the study of sound change, speech perception and production. In the model, the driving force for sound change is an error in perception caused by production of co-articulation. In short, "universal and timeless physical constraints on speech production and perception leads listeners to misapprehend the speech signal. Any such misapprehension that leads the listener to pronounce things in a different way is potentially the beginning of a sound change." To explain this in more detail, I will describe a current experiment done by Kleber et al. (2010), which provides further evidence for the validity of Ohala's model.

There were 33 Standard English speakers (assigned to either a young or an old group) participating in this experiment. It was divided in two parts: one investigating speech production, the other investigating perception. In the first part, participants produced various syllables containing /ʊ/ in different consonantal contexts. Results indicated that the young group produces /ʊ/ in the (s_t) context fronter than in the (w_l) context. This effect is weaker in the old group. In the second part, they were presented with syllables containing a vowel from the /ɪ-ʊ/ continuum embedded in either (s_t) or (w_l) contexts, which are known to induce vowel fronting and backing respectively. A forced choice had to be made with regard to the vowel being /ɪ/ or /ʊ/. Results showed that both groups categorised the vowel /ʊ/ more often in the (s_t) than in the (w_l) context. Additionally, the young group inclined significantly more towards /ʊ/ than the old group. This proves that /ʊ/-fronting is a vowel shift in progress in English. The young generation grew up hearing different /ʊ/'s in different contexts and recognised an intention in the speaker of these utterances to pronounce them differently. When they take over this intention, the differences are amplified even more, resulting /ʊ/ in (s_t) to be even fronter.

Chapter 3

AI models of language evolution

Since the dawn of artificial intelligence in the 1950s, the field has slowly but surely infiltrated neighbouring fields. Offering new methods, data and insights, this infiltration gave rise to a boost in these respective fields. In the early 1990s, it was the turn for the field of language evolution, and the pioneer infiltrator was Luc Steels. His Talking Heads experiment (Steels, 1999) was the starting point for a series of thought-provoking simulations and experiments, based mostly on language games and iterative learning.

Three of many good reasons for simulating language evolution are mentioned in Cangelosi and Parisi (2002). First, by making computational models of theories, ensures that theories can only be explicit. Often you find that a researcher has an idea of what is happening and posits a theory of a still sketchy nature. By simulating it, you ask of the researcher to become concrete and not cut corners: if the theory is not explicit, the simulation will not work. Second, simulations generate a vast amount of detailed predictions, that are otherwise difficult to attain. When the system being investigated is very complex, it is often hard to tell what happens under certain circumstances, resulting in speculation or guessing. With a computational model it becomes possible to see the resulting behaviour, no matter the system's complexity. Third, there is the possibility of manipulating variables in the setting of a virtual lab. Once the model is complete, you can test hypotheses easily by running a simulation twice with different parameter settings.

After two decades of cooperation, there is still a communicative gap between the linguist and the computationalist. Their completely different education have rendered them speaking two distinct languages. The one talks about programming and designing, the other talks about producing and comprehending. To tackle this problematic situation, one of the goals of the Evolang workshop was to increase awareness in computational modellers of how models fit together (de Boer and Zuidema, 2010). Modellers are advised to be clearer and more open in describing what kind of model they use, what it can do, and how it and its assumptions relate to other models and reality. Perhaps in the future, there will be centralised information that bonds the two fields of linguistics and artificial intelligence.

Gong (2010) lays down 4 critical steps that need to be taken when creating a computational model of language evolution. I wish that I had known about these guidelines before I started my MSc project, because in retrospect they correspond with the steps I eventually took. Besides, they are just very clear. These are the steps:

1. set up an artificial language,
2. define linguistic knowledge and learning mechanisms,
3. implement a communication scenario, and
4. analyse the system performance.

I will refer to these 4 steps in the descriptions of both the simulations from the literature as the one I implemented myself.

The language game is often seen as a solution to the third of the 4 steps. In general, language games are interaction protocols for agents in simulations. They were introduced by Steels in the mid 1990s, himself being inspired by Wittgenstein and bird song (de Boer, 2001, p.51). These protocols prescribe when agents have to perform which action and what updating processes need to be run when the game is over. Of course these agents need to be able to comply to this protocol, otherwise they are not fit for the job. Examples of often used language games are the guessing game, seen in the next section about the Talking Heads experiment, and the naming game. In Chapter 5, a new language game called the aligning game is introduced. This game is used for the simulation in this thesis. In this chapter, the pioneering work is described in considerable detail first, after which recent efforts to simulate vowel systems and linguistic change pass in review.

3.1 Talking Heads experiment

The Talking Heads experiment (Steels, 1999) acts on the level of communities and falls under glossogeny. There was a set of installations at different places in the world which contained physical bodies for the agents and an environment or scene which the agents could communicate about. The agents' architecture exists of five distinct layers, within which an hierarchical structure exists. A

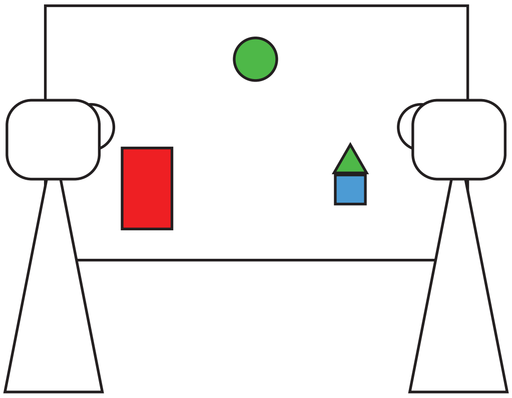

Figure 3.1: An example environment in the Talking Heads experiment. Two agents look at a white board with their cameras. The white board environment contains several objects about which the agents can communicate.

summary of these layers is shown in Table 3.1. I will explain each layer using the example scene in Figure 3.1.

Suppose an agent wants to give a hint about the green object in this scene. It does that by gathering data from image segments via its camera, such as Size = 5 and Color = #00FF00. It then categorises the green object's segment as being small and green, based on a self-generated repertoire of categorical distinctions called the ontology. The agent now retrieves words it has linked to these distinctions from another repertoire: its lexicon. Say that it associates small with the word "nuguge" and green with "xani". The fourth layer allows the agent to combine these terms to form a question like "nuguge xani?". As you can see, each layer represents one step in the semiotic square (Figure 2.1 on page 6), either forwards or backwards. This implementation completes the first of the 4 steps.

Name	Function	Output
Perceptive layer	Break down image into segments	Size = 5
Conceptual layer	Categorise segments into distinctions	Small
Lexical layer	Associate distinctions with syllables	'nuguge'
Syntactic layer	Organise syllables into larger structures	'nuguge xani?'
Pragmatic layer	Runs guessing game script	—

Table 3.1: The Talking Heads' architecture in five distinct layers. For each layer, its name, its function and an example output form is given.

When two agents are loaded into robot bodies in the same location, they both become Talking Heads and are eligible to play a guessing game. In this language game, there are two agents, a speaker and a hearer. First, the speaker chooses an object from the environment called the topic, and gives a verbal hint to the hearer as to identify this topic. It chooses a verbal hint based on previous successfulness of hints. After this, the hearer tries to guess the topic that the speaker has chosen by pointing to the object of choice. The speaker sees where the hearer points and indicates if this guess is correct or not. If the hearer did not understand the hint in the first place, for instance because the uttered word is not in its lexicon or because there are no objects in the context that fit the description of the hint, it asks what object the speaker meant and remembers the meaning of the word according to this.

After a language game, adjustments are made in the participating agents' lexicon and ontology. If the game was successful, the speaker increases the value of the word-meaning pair used to give a verbal hint with δ and decreases the value of other words with the same meaning with the same δ. The latter is a clear form of lateral inhibition and makes sure that any competition between synonyms will be won eventually. The hearer increases the value of the word-meaning pair with δ as well, but instead it decreases the value of other meanings of the words. These mechanisms complete the second of the 4 steps.

Luc Steels wanted to test several hypotheses with this experiment, four to be precise. These hypotheses are much in line with the view of languages as a CAS. First and importantly, Steels puts forward the idea that "language emerges through self-organization out of local interactions of language users. It spontaneously becomes more complex to increase reliability and optimise transmission across generations of users, without a central designer." (Steels, 1999). What he means is that language is a joint effort of the sum of the parts of a system, here a network of agents. There is no predefined authority that invents or manages it. Results of the experiment favoured this idea, for it was clearly seen that a shared lexicon arose after a period of competition between synonyms and ambiguities.

Second, Steels hypothesises that meaning is built up slowly by the individual. Meaning is not innate or learned through induction but rather built up with experience. This experience comes down mostly to the interaction (guessing games) with other agents. The successfulness of these games gives an idea of how successful categorical distinctions used as meaning are, thus making language and meaning co-evolve in the Talking Heads experiments.

The third and fourth hypotheses are of little importance for this thesis. They concern the notion that an ecology is a good metaphor for a cognitive system and the spontaneous development of grammar. So overall, the Talking Heads experiment made use of an elegant bottom-up approach to simulate the evolution of a shared lexicon. It succeeded in showing that language can very well

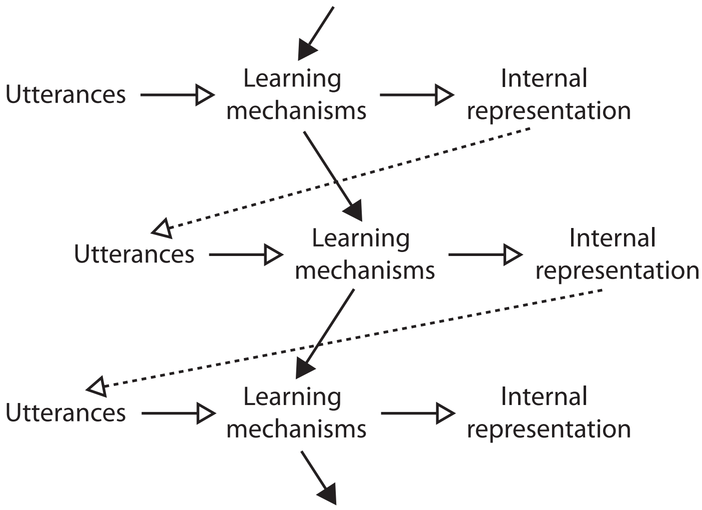

Figure 3.2: Evolutionary iterated learning in a nutshell: Learning mechanisms are passed on through biological evolution, while they allow for internal representations to be formed based on the perception of utterances. These internal representations themselves are at the basis of new utterances.

emerge through local interactions between users. It also jumpstarted a large amount of simulation research on language evolution related topics. Vogt (2003) has even made a software simulation version of the experiment.

3.2 Edinburgh experiments

Simon Kirby, colleagues and students at the University of Edinburgh have been laying out a new framework for the modelling of language evolution. Especially over the last decades, a number of stunning experiments were conducted by this group, revealing that a structured language evolves in the laboratory without intentional design of participants. All simulations and experiments are based on the iterated learning model (ILM). Below is a short description of the ILM and a recent experiment that makes use of it.

The ILM is a so called Expression/Induction (E/I) type model (Hurford, 2002). What characterises E/I models is the identification of two media through which language evolves, namely the physical utterances made by speakers (Chomskyan E-language) and the mental representation of them within the speakers (I-language). This corresponds surprisingly well with the division made in Figure 2.1 between the external and internal parts of communication: language changes both internally, for instance in the speaker's lexicon, and externally in the signals produced. It is inherently impossible for an individual to directly alter the internal representation of another individual, instead this happens via the external world by communication.

What characterises the ILM in specific is that it is a framework in which the output of one becomes the input of another person. Typically, an adult agent produces an internal set of signal-meaning associations (I-language). This is used to generate utterances (E-language) from which a new generation learns its own I-language. These agents become adults and the script continues. A visual model of evolutionary iterated learning can be seen in Figure 3.2. Here, besides the transmission of I-languages through utterances, it can be seen that the process of biological evolution of the learning mechanism is also present. The ILM thus provides an implementation for the second of the 4 steps: defining learning mechanisms. The third step is generally much less emphasised.

Kirby et al. (2008) apply the ILM in a recent experiment with actual participants. This contrasts with the usual software agents, turning the usual virtual lab into a real one! In the experiment, participants are asked to learn an alien language, which exists of orthographic signals (labels) associated with visual stimuli (images). Images differed along three dimensions, being shape (square, circle and triangle), colour (black, blue and red) and motion (straight, bouncing and spiraling).

Initially, random labels are generated for all animations, thus constructing an unstructured language, unless by chance the random generation did a bad job. Participants learned half of the 27 label-image pairs in a learning phase. In the next phase of the experiment, they were asked to create labels for all images. The output of one participant was then the input of the following participant, according to the ILM. The main hypothesis was that cumulative adaptive language evolution would be seen. In other words, what was first a random language is expected to become more structured and learnable over the course of participant generations. What they found is that with a chain of 10 generations, this was indeed the case. Whereas the first generation's language was still random, the last generation had often converted it to a structured language, in which particular features of the label word came to represent features of the image. For example, in one resulting language, black objects started with an 'n' and blue ones with an 'l'.

3.3 de Boer's vowel systems

Before calling the Backgrounds chapter quit, I have to describe another refreshing and creative use of simulation, namely that of the Ph.D. work done by de Boer under Steels' supervision (de Boer, 2001). The aim of this simulation was to research vowel system emergence in a community of software agents that learn to imitate each other. The agents were modelled to be as much like humans as computational costs would allow. They were able to both produce utterances of single vowels through an articulatory model, consisting of the three parameters position, height and rounding. Based on values for these parameters and a set of synthesizer equations, the first four formants of the acoustic signal could be computed and a sound could be produced. Both (uniform) articulatory and acoustic noise were added to the mix of this production process. Agents could then perceive this with an perception model. This is also very human-like, since agents perceive the vowels as prototypical vowels, instead of simply the formant frequencies. They have a vowel set (much like a lexicon but then without meaning linked to it) to which each incoming speech sound is compared. The prototype vowel in the set closest to the perceived signal is considered to be actually heard.

Agents engage in imitation games, another form of language games. An imitation game takes two agents, an initiator and an imitator. The initiator chooses a vowel from its prototype vowel set and produces it, with added noise. The imitator perceives this, hears one of its prototype vowels and produces it, again with added noise. Then the initiator perceives and hears this. If what he hears is the same as he chose, the game is successful, otherwise it is not. The initiator communicates the outcome non-linguistically to conclude the interaction part of the imitation game.

Then follows the updating. In case of success, the imitator shifts the used vowel to the one perceived in acoustic space. It does that by considering the acoustic output that its six neighbours in articulatory space would give. These neigbours each differ from the original prototype in one parameter by a small amount, either positive or negative. It then keeps the one which most closely resembles the perceived signal. In case of failure, two things can happen. If the vowel has been successful in the past, the agent creates a new vowel in the middle of the vowel space. If it hasn't, the agent shifts it as described above. Additionally to the imitation games, the agents' vowel sets were modified by clean-ups and mergers, which occurred randomly.

As for the results, first of all the agents succeeded in forming a shared vowel system. Depending on the amount of noise, systems with different numbers of vowels emerged. These resulting vowel systems were compared with what Crothers' typology (see Figure 2.3 on page 8) would predict the systems to be like. Most of the systems lived up to these predictions. For example, 78% of the three-vowel systems had /i a u/. Within four vowel systems, all had /i a u/ and 55% had /ɘ/ and 45% had /e/. For larger systems, vowels were added that are quite consistent with Crothers' typology. I can't help but observe that /ɘ/ comes into play relatively prematurely. Perhaps the cause of this is that new vowels are always added in the middle of the vowel space, which is exactly where /ɘ/ is located. Still, this indicates that self-organisation is enough for agents to develop realistic vowel systems, which is quite an achievement.

3.4 Recent simulations of linguistic change

There is other great work being done. Please go and read about it, I can only mention some of it here.

Livingstone and Fyfe (2000) simulated the emergence of dialects in a small population of agents, which were placed in a single row without the ability to move (see also Livingstone (2002)). Agents could communicate with other agents within their neighbourhood, based on a distance parameter. The results of this simulation showed the formation of dialect continua, whereby agents that lived close by could understand each other, while agents farther away could not. When the distance parameter was increased, the continua converged into one shared language. Interestingly, adding acoustic noise prevented this convergence.

Both Dowell (2006) and Hira (2009) did simulation experiments on the subject of Creole languages for their master theses. Dowell based his simulation on de Boer's work and expanded it by letting different shared vowel systems come together to form a Pidgin or Creole language. Based on literature studies, pidgin languages were expected to have a lower amount of vowels than creole languages, but the simulation results proved the exact opposite. Dowell ascribed this lack of similarity with reality to the simplicity of the model. Hira found that when two communities with different languages come together, they are often able to end up with a common language using language games. Also, two communities with the same initial languages were shown to either stay identical, become partly different or completely different, depending on the time they spent together in the first place. When they spent much time, no new words were created and the languages remained the same.

Kwisthout et al. (2008) investigated the role of joint attention in language evolution. Each of the three stages of joint attention identified in children's early development (checking attention, following attention, and directing attention) was modelled computationally. This resulted in eight type of language games, in which the three types of attention were either absent or present. The best results were found in simulations where checking attention was present, the second best results in those involving following attention. This was in accordance with the order in which joint attention is acquired in young children. It was also argued to be the order in which joint attention emerged in human evolution. Schepens et al. (2010) sought for cognates in a professional translation database of 6 European languages. Translation pairs in this database were categorised as cognates if both semantic and orthographic overlap was high. Resulting cognate distributions were found to be similar to language similarity orderings based on both (Gray and Atkinson, 2003) and the outcome of a language similarity validation questionnaire.

Chapter 4

The linguistic dynamics model

In this study, evidence is provided for a model of linguistic change. This model is visualised in Figure 4.1. Most of its elements are not new; what is new about the model is the synergy of its elements. Compact as it is, this model is able to explain a wide variety of phenomena. Three hypothesis follow from the model, namely:

Formation hypothesis reinforcement of communicative success and punishment of disuse drive the formation of a shared, intelligible lexicon,

Change hypothesis alignment to variation in perceived signals drives linguistic change, and

Divergence hypothesis social fragmentation caused by geographical constraints drives linguistic divergence.

To assure that these hypotheses are correct, they are put to the test. In chapters 6, 7 and 8, the respective hypotheses are backed up by simulation evidence. The software tool used to run these simulations is described in Chapter 5. It involves the aligning game, a new type of language game based on the proposed model. In this chapter, a description is given of how the model came about, what it says, and how it should be tested.

4.1 Model foundations

The formation hypothesis is inspired by the idea that language emerges through self-organisation out of local interactions of language users (Steels, 1999) and that local interactions (utterances) are the basic elements of language (Croft, 2000). The actual content of the hypothesis arose from a feeling of discontent about the processes currently believed to enable the formation of a shared lexicon. Let

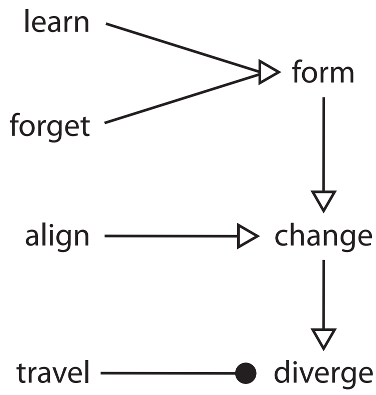

Figure 4.1: Diagram of the model of linguistic dynamics. Arrows indicate that when the former increases, the latter increases as well. Lines ending in a filled dot indicate that when the former increases, the latter decreases.

me recreate this feeling in the reader and explain how the presently tested hypothesis emerged from it.

Lexical formation is the transition in a community of being without language to having an intelligible language. According to the idea of self-organisation, this transition is made possible by processes that alter the individual lexicon on the basis of local interactions. This type of phenomenon, in which complex and global behaviour comes forth from simple local behaviour, is often called emergence. The Talking Heads experiment provided considerable evidence for this view. In this experiment, software agents (models of human speakers) were placed together, initially without any language. Over the course of a simulation of local interactions, the agents formed a language.

How did the agents manage this? Well, the updating processes that altered the individual lexicon in the Talking Heads experiment come down to the following five:

1. creating an association when none is available,
2. taking over an association from partner when it is unknown,
3. rewarding an association when it is used successfully,
4. punishing an association when it is used unsuccessfully, and
5. punishing an association when a synonym or homonym of it is used successfully.

Let us briefly go through these processes.

The first process seems to be a necessity. Without ever creating associations between a signal and a meaning, the agent will remain languageless. When you have no word to describe some meaning, why not make one up? Whether the action is as easily performed by a human as it is by a software agent remains a question to be answered; I will not discuss this presently.

The second process entails another useful necessity. When the person you speak with uses a word you do not know, it is only natural to learn this association by taking it over from this person. These two actions round off what one could call the creative processes. With these processes, a human is able to acquire a language. The other processes adjust the strength of the created associations.

The third process allows associations to become more rooted in the person's mind. Whenever a word is used successfully, it gains power, thereby making it more likely to be used again, thereby—once more—increasing the language's intelligibility. This process is the 'learn' node in Figure 4.1, although one could count the first two processes to be included there as well.

Although the previous process is plausible to exist in humans and their brains (neuronal paths being strengthened by being used), the fourth process strikes me as implausible and artificial. Is an association weakened when it is used unsuccessfully? Perhaps the fact that it is being used at all is reason for it to be strengthened instead of weakened. For it to weaken would require a higher mechanism perhaps controlled by the recognition of failure that comes after the use.

The fifth process is even more artificial. Imagine you are Dutch and someone is talking to you about sitting on a bank. In Dutch, the word 'bank' can denote a monetary institution, as in English, but also a piece of furniture: a couch. According to the synonym punishment process, this interaction causes you to weaken several associations in your brain, under which that of bank with monetary institution and that of sofa with piece of furniture. This is an unrealistic, farfetched and complex process. It is not only a computationally heavy procedure in the simulations: it would require an ingenious system to explicitly implement the results of this mechanism in the brain as well. Yet, Steels (1999, pp. 142–143) holds synonym punishment to be crucial for lexical formation.

In this thesis, I argue that punishment of disuse is a better alternative to these two unrealistic punishment processes. Punishment of disuse entails simply that when you do not use an association, you gradually forget it. This memory decay acts as quicksand where associations can only survive by being used successfully. Besides being a less selective mechanism, it allows for synonymy and homonomy, which is realistic because this is often seen in human speech.

Punishment of disuse is represented in Figure 4.1 by the 'forget' node, although it could just as well have been called the 'decay' or 'disuse' node. If this process does a similar job in allowing a community to form a shared language, it should be preferred over the other punishment processes, because of its plausibility and its simplicity. Ockham's razor says that when two theories compete, the simplest explanation is to be preferred. Because the idea of punishment of disuse makes almost no assumptions, is simpler and permits the existence of synonyms, it should, according to Ockham's razor, be preferred.

The change hypothesis came about as follows. In recent simulations on lexical formation, the ultimate goal is to have a population of agents to develop a shared language. Although this sounds like a worthy goal, stability and homogeneity is not at all what is observed in reality. Languages change constantly on a personal and communal level. Yet, the models used to simulate lexical formation are unable to show this change and variance. Hira and Bartlett (2010) also note this absence of linguistic variation. Steels (1999, p. 238) himself also mentions the fact that languages as a whole shift in the real world.

When you actually look a bit deeper at the assumptions behind these simulations, this is not strange. None of the five updating processes mentioned before actually alter the nature of the signal. They only create, strengthen, or weaken it. Furthermore, new associations are only created in the context of unsuccessful games, indicating that when a language settles and all games are successful, variation is reduced to zero. In order to account for linguistic diversity and change, one must pinpoint the mechanisms by which signals change in reality, model them and test them.

I propose alignment, in which agents adjust their pronunciation to their partner's pronunciation, to be such a mechanism. This updating process exploits existing or agent generated pronunciation variance to enable linguistic change, given that the adjustment is to some extent permanent. Without a process like alignment, variation remains largely in the external language, in the signals used. Alignment can be seen as welcoming this variation to the internal language, to the minds of the users and/or their lexicons. With alignment, the hearer adjusts its internal signal towards perceived signals, just as HRH Queen Elizabeth II may have done.

Pickering and Garrod (2004) and Pickering's (2006) studies about alignment boosted my confidence in the hypothesis that language change may arise at least in part from alignment. Different rationales have also been reported for agents to align, such as maximising communicational success, gaining social acceptance and communicative grooming (Giles and Ogay, 2006). Yet, remarkably, the view of alignment as a source of linguistic change has not been adopted often, or at all. Regarding alignment studies, emphasis is put on the reasons why people align, and the short-term effects of alignment, not the long-term ones like linguistic change.

The divergence hypothesis is inspired by the observation that there is a diversity of languages in the world. Of course, when there is no linguistic change, no difference can ever arise from a community that initially has a shared language. Here, this thesis' simulations hopefully differ from standard lexical formation studies. Given that linguistic change occurs in our simulations and that it is driven by individual interactions, I reasoned that which agent interacts with which agent decides where change occurs.

In order to arrive at a system of multiple dynamic languages as we see in the real world, one needs to bear in mind that linguistic change only occurs under the presence of communication. If everyone talks with everyone, one shared language will ideally be formed. However, when the population is somehow divided into groups such that communication is present within groups but absent between them (in other words, when there is social stratification), linguistic change will occur in each group, independent of other groups. A logical reason for such a division to arise are geographical constraints: when there is an impassable ocean between two communities, they are divided and contact becomes impossible. In this case, it can only be such that these communities' languages change in their own direction, ultimately creating linguistic divergence.

4.2 How to test the model

The model will be tested using the simulation tool described in Chapter 5. In this tool, agents link vowels to referents to create a language. With it, the model of linguistic dynamics can be tested for the specific domain of phonology, more specifically that of vowels.

In the Chapter 6, six agents play language games with varying amounts of reinforcement and decay. The prediction is that lexical formation only occurs under the presence of reinforcement and decay. Proving this prediction is true would be a leap forward in understanding linguistic dynamics in general. To reiterate, the fourth and fifth process have been replaced by a new process. Another process will be added to test the second hypothesis.

In the Chapter 7, six agents play language games with different aligning rates. What this hypothesis predicts is that without alignment the language remains static, and that with increasing amounts of alignment, languages change with increasing speed. Especially agent generated variance is expected to act as a setter of new targets for the alignment process to aim for.

This hypothesis is tested in the Chapter 8, wherein 12 agents with the same initial language are divided over two locations. Geographical constraints between two locations decrease the rate of travel between these locations. The agents play language games and travel between locations with varying travel rates. The prediction is that absence of travelling leads to the two languages diverging, and increasing amount of travelling causes this divergence to slow down. This would support the divergence hypothesis, because absence of travelling occurs under severe geographical constraints, and should therefore induce divergence.

Chapter 5

The simulation tool

The simulation tool I made during my internship is called the Dialect Emergence Virtual Lab (DEViL). The DEViL falls under glossogeny and acts on the community level. de Boer and Zuidema (2010) distinguish models of linguistic evolution based on function, form and validation. The has an explanatory function, as it models only the essentials needed to explain linguistic dynamics. The DEViL takes the form of a computational model, as opposed to a mathematical one. As we will see in the following chapters though, it is at times still possible to describe the resulting behaviour formally. As for the validation of the model, due to the limited time one has for a Master project, in the simulation chapters validation is limited to internal validation. Note that externally validating vowel change and divergence in a formal way is not straightforward at all. If you have read the Principles of linguistic change section, you will understand that implementing algorithms that test if all principles of vowel chain shifts are abided by in the simulation population is indeed quite a work. Besides, given the abstractions made away from realistic perception and production, I don't expect them to: remember that the goal is to check if linguistic dynamics arises from certain parameters, not if realistic dynamics arises. This means that it is only tested if the predicted behaviour follows from the model hypotheses. External validation is moved to the General discussion section, where it is done more informally.

The DEViL is programmed in Java. All code is written by myself, although I made use of the JFreeChart API and the Batik SVG toolkit to make a bubble plot of the signal space and to save these as vector images respectively, so thanks to the creators and moderators of JFreeChart and Batik. Parts of the code are available for viewing directly in Appendix A. To visualise the results, I made a plotting function in Matlab. In short, the tool enables the user to created worlds that exist of locations connected to each other through paths to form a map, or a network. These locations hold different kind of objects, called referents, which agents can interact about. A population of agents can be inserted in this world, either with no language at all or with a predefined one. Furthermore, the user can set simulation parameters and then run the simulation. It is possible to actively view the communal signal space evolve in real time. The tool makes generates several output files, containing information such as the state of the system, measures over time and images of the signal space. In this section, I will describe the essentials.

5.1 The agent

The architecture of our agent is a stripped down version of the Talking Head. Whereas the Talking Head associates signals to referents through observation and meaning, our agent makes a direct association between referent and signal. This abstraction has been made to make it faster for me to implement the tool. Meaning is also not a prime target of the aims of this thesis, making it a suitable element to abstract away. As for the signal space, I have chosen for the phonetic domain: the vowel space, just as de Boer (2001) did. Agents thus link referents to vowels in their lexicon. These vowels consist of the first two formants, which both take real values between 0 and 100. The values to the left of the line $f_2 = 100 - 2f_1$ and to the right of the line $f_2 = 2f_1 + 100$ are off limits, to create a vowel triangle spanned up by the vertices [0,100], [50,0] and [100,100]. Through linear transformation, these numbers could be altered to become realistic frequencies or bark scale figures. There are four reasons why I chose to merge perception and production into one frequency based lexicon. The first is speed of computation, the second is speed of implementation, the third is simplicity and the fourth is that alignment is not a mechanism that necessarily requires the implementation to be divided into perception and production. In the future I would like to make a more realistic model, since other mechanisms such as co-articulation do seem to rely more on this division.

The agents in the simulation can perform two general actions: walking and talking. Regarding the first, every time step the agent decides if it will travel. It then decides its destination location and travels there. Talking is explained in detail in the following section.

5.2 Aligning game

I have designed and implemented a new type of language game, which is based roughly on Steels' guessing game and the model of linguistic dynamics proposed in the Introduction. In the description of the model, five processes emerged that change the agent's lexicon. So you don't have to leaf back, these were:

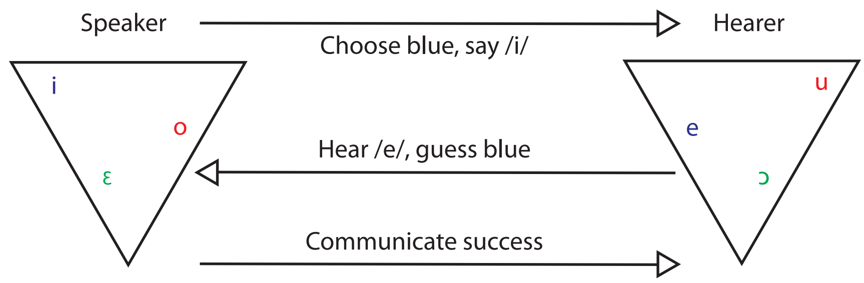

Figure 5.1: An example aligning game, picture setup inspired by de Boer (2001).

1. creating an association when none is available,
2. taking over an association from partner when it is unknown,
3. rewarding an association when it is used successfully,
4. punishing an association when it is not used and
5. adjusting a signal towards partner when used successfully.

These are all incorporated in the new game, which is called the aligning game, since alignment is a new process in language games and thereby a distinguishing feature of this particular game. Next, I describe the aligning game, indicating where the five processes are located. In Figure 5.1, an example aligning game is showed.

Just as in the classic guessing game, the aligning game is played by two agents, a speaker and a hearer. The speaker chooses a referent in the location and attempts to communicate this to the hearer by giving a hint. Two things can happen: either the speaker has no vowel associated with the referent, in which case it creates one (process 1) and signals this new vowel to the hearer, or the speaker does have one or more vowels associated with the referent. In that case it picks a vowel that denotes the referent, each vowel having a chance of being chosen that is proportional with its associative strength (so strong associations have a higher change of being picked), and signals that vowel to the hearer.

The hearer perceives the signal and attempts to guess which referent the hint was intended to label. It does this by simply retrieving the association which vowel is closest in the vowel space to the hint. To communicate the guess, the hearer points to the referent of this retrieved association. If the hearer has no language yet, it just shrugs its shoulders.

The speaker decides upon the success of this aligning game by comparing the referent it chose with the referent the hearer pointed to. If this is the same referent, it will increase the associative strength of the used association (process 3). Otherwise, it does nothing. Either way, at this moment the speaker informs the hearer about the game success by pointing at the intended referent.

If the game was successful, the hearer rewards the association it used (process 3) and moves the vowel of this association in the direction of the hint in vowel space (process 5). If the game was not successful, the hearer takes over the hint as a new association with the intended referent (process 2).

If you paid attention to the indicated processes, you see that process 4 is missing. This is because it is not dependent on interaction. On the contrary, punishment of disuse acts as a function of time, slowly decreasing the associative strength of all associations. Thus it creates a quicksand, which the associations can only keep there figurative heads above by being used in successful aligning games.

5.3 Parameters

You can set a lot of parameters in the DEViL.

The map on which agents are located can have different number of locations, which is controlled by L. R sets the number of referents present in each location. The number of agents can be varied with A. By default, all locations are connected to each other with the highest connectivity through having geographical weights ω_g of 1.0. To further customise the simulation environment, you can edit these ω_g in the .map-file. By setting it to 0.0, traveling becomes along that path becomes impossible. These weights are an attempt of a numerical representation of the differences in traveling difficulty one encounters in the real world, for instance due to seas, mountains and other things influencing free trespassing.

Another way of controlling the amount of traveling is the travel rate τ, which is defined as the probability of an agent deciding to travel each time step. τ should be set somewhere between 0.0 and 1.0; agents travel always when τ is 1.0, they never travel when it is 0.0. If an agent decides to travel, the chance it travels along one particular path is proportional with the ω_g of that path, making it more likely for agents to travel paths with high geographical weights. Furthermore, in the .map-file you can also edit which referents are present in which location and in the .pop-file you can assign locations to each agent and even add lexical associations to agents.

Ψ_{ac} sets the amount of production noise added to the vowel when it is transferred from speaker to hearer. It is implemented as Gaussian noise with a σ^2 of the value set for Ψ_{ac}, see Figure 5.2. This parameter models differences between intended and perceived pronunciation by computationally approximating the effects of different processes going on in the time span from forming intention and perception. These processes might include differences in physical condition of the agent, differences in positioning of the speech organs, errors in perception and background noise interfering the sound transfer. For now, I grouped these into one parameter; in the future it might be interesting to dig

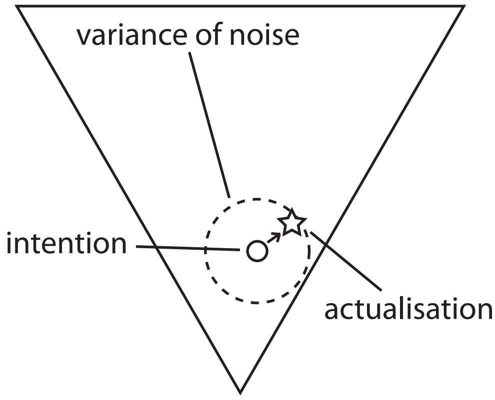

Figure 5.2: An explanation of the implementation of acoustic noise Ψ_{ac}, picture setup inspired by de Boer (2001).

further into this, by modelling them separately.

δ_r governs the rate of reinforcement after a successful aligning game. It can be set to any positive number and simply states how much the associative strength ω_α is increased. For example, when an association has an ω_α of 5.4 and δ_r is set to 0.2, ω_α will be 5.6 after a successful game. δ_d governs the rate of memory decay, or punishment of disuse. Every time step, all associations are decreased by the value set for this parameter. This means it is typically set as a small positive value. δ_a governs the rate of alignment. After a successful game, the hearer moves the vowel it used towards the hint with the amount set for this parameter. ω_0 is the associative strength an association starts out with. By default this is set to 10. ω_{max} is the maximal strength an association can have, which is set to 100 by default. ω_{min} is the minimal strength of any association, which is always 0. So the default association strength space spans from 0 to 100.

5.4 Measures

The DEViL offers various measures that can be used to track the state of the simulation. I describe them in this section.

A first measure is synonymy (s), which is defined as the average amount of vowels a referent is associated with. It is computed as

$$s = \frac{\sum_{a \subset A} \sum_{\alpha \subset \aleph_a} 1}{RA}, \tag{5.1}$$

in which $\aleph_a$ is the total set of associations in agent a's lexicon. If each referent is associated with one vowel, as in an optimal language, then $s = 1$.

Another measure is average association strength ($\bar{\omega}_\alpha$). It is computed as

$$\bar{\omega}_\alpha = \frac{\sum_{a \subset A} \sum_{\alpha \subset \aleph_a} \omega_\alpha}{As}, \tag{5.2}$$

in which ω_α is the association strength of association α. Just as ω_α, $\bar{\omega}_\alpha$ is bound by ω_{min} and ω_{max}.

Communicative success (c) is the next measure. It is computed by keeping track of the number of successful games and the total number of games. Dividing the first by the second gives you c.

Three measures are related to lexical distance (LD). Lexical distance measures aim to quantify how much languages differ. The aim for these measures is for identical languages to have an LD of zero, and of increasingly different languages to have increasing values of LD.

Local variance (LV) is one of these measures. It measures the amount of variance in the lexicons of local inhabitants. If all agents in a location have exactly the same language, LV is 0; with increasing variance, LV increases. The computation of LV is more involved: first we compute cumulative lexicons $\aleph_l$ for each location l with

$$\aleph_l = \frac{\sum_{a \subset l} \aleph_a}{\sum_{a \subset l} 1}, \tag{5.3}$$

in which $\aleph_a$ is the lexicon of the agent. We then define vowel distance VD between two associations' vowels as

$$VD(\alpha_1, \alpha_2) = \sqrt{(f_{1,1} - f_{1,2})^2 + (f_{2,1} - f_{2,2})^2}. \tag{5.4}$$

Note that this distance is Euclidean. In the equation, $f_{1,2}$ is the first formant of the second association, $f_{2,1}$ is the second formant of the first association, and so forth. Then we define average association $\bar{\alpha}_r$ within a lexicon for a certain referent r as

$$\bar{\alpha}_r = \frac{\sum_{\alpha \subset \aleph_r} \omega_\alpha \alpha}{\sum_{\alpha \subset \aleph_r} \omega_\alpha}. \tag{5.5}$$

Note that each association is weighted according to its associative strength. Then, finally, LV is

$$LV = \frac{\sum_{a \subset l} \sum_{r \subset R} VD(\bar{\alpha}_r(\aleph_a), \bar{\alpha}_r(\aleph_l))^2}{R \sum_{a \subset l} 1}, \tag{5.6}$$

in which $\bar{\alpha}_r(\aleph_a)$ is the average signal for a referent in an agent's lexicon, and $\bar{\alpha}_r(\aleph_l)$ is the average signal in the local cumulative lexicon.

Two other measures of lexical distance are historical lexical distance (HLD) and interlocal lexical distance (ILD). HLD compares a language with its ancestor language, while ILD compares languages of different locations. Both of these measures are comprised of three parts: a vowel space based distance, a synonymy based distance and an associative strength based distance. These parts are mixed in a 3:1:1 ratio to form both HLD and ILD. Since the computation of each of these three parts is at least as involved as the computation of LV, I refer to the Appendix for the implementation of these measures.

Two measures are related to lexical formation: formation success and formation speed. Both are usually measured over several simulations. In general, a shared lexicon is defined to have formed when $c > .75$ and $LV < 5$. The $c > .75$ condition covers successfulness, whereas the $LV < 5$ covers sharedness. Formation success is defined as the percentage of simulations which achieved lexical formation within 10,000 time steps (or another number of time steps, but this will always be stated). Formation speed is defined as the average time step at which successful simulations achieved lexical formation.

Chapter 6

Lexical formation

The main goal of this chapter is to test the hypothesis that reinforcement of successful communication and punishment of disuse are the processes that allow us to form a shared, mutually intelligible language. I do this by simulating lexical formation with the DEViL tool. The parameters δ_r and δ_d represent reinforcement of successful communication and punishment of disuse respectively.

6.1 Variables

The simulation is set up as follows. In the beginning, 6 agents are placed together in one location. These agents have no initial lexical associations. In other words, they are without language. The location holds 6 referents that the agents can interact about. The agents will play aligning games (without actually aligning just yet, this will be introduced in the next chapter about vowel change) as explained in the Methods section. They do so until 10,000 time steps have elapsed, or until the agents' lexicons become too large for the simulation to finish in a reasonable amount of time (this threshold is set at a synonymy of 50).

All parameters except reward and decay rate are kept constant throughout this chapter. Acoustic noise ψ_{ac} is set to 5, travel parameters ω_g and τ are irrelevant since there is only one location, the ω_{min}, ω_0 and ω_{max} are set to the default values of 0, 10 and 100 respectively. There are two (independent) variables in this simulation: rate of reward δ_r and decay δ_d. δ_r takes values 0, 0.5, 1.0, 2.0, 4.0, 7.0, 10.0 and 15.0, while δ_d is set on either 0, 0.05, 0.1, 0.2, 0.4, 0.7, 1.0 and 1.5. Together, these settings make up 64 different simulations. Each of these is run hundredfold.

Three measures are used in this chapter to function as dependent variables: formative success, speed of formation and synonymy. First, the formative success is the percentage of cases in which the population successfully formed a shared vowel space. The second measure is the speed of lexical formation, that is the average amount of time elapsed before the lexical formation stop condition is reached. Third, synonymy is the average number of vowels associated with a referent. A more detailed description of these measures can be found in the measures section.

6.2 Prediction

The prediction is that lexical formation only occurs under the presence of reinforcement (reward rates higher than zero) and punishment of disuse (decay rates higher than zero). In stark contrast to the Talking Heads experiment, synonym punishment is not assumed to be necessary for lexical formation. Neither is punishment of associations used in unsuccessful interactions assumed to be needed. Instead, the predictions are based on the first hypothesis of the model described in the introduction in which only reinforcement and punishment of disuse are the driving forces behind lexical formation. The hypothesis is thus that formative success is 1 in all cases where both reward and decay rates are higher than zero and 0 in all other cases. The speed of formation can give valuable extra information about the formation process and how it is affected by reward and decay rate, as does the measure of synonymy.

6.3 Results

The results of the proposed simulation can be studied in Table 6.1 and Table 6.2. As can be seen, the prediction has come true insofar that lexical formation does not occur when either memory decay or learning from success is absent: in all cases where δ_r or δ_d is 0, 0% of the simulation runs resulted in a shared lexicon. The other half of the prediction was that formative success is 100% in all other cases. This appeared not to be the case. For low values of δ_r, formative success is still 0%. Only for values higher than 1.0, shared lexicons start to form. For higher values of δ_d, lexical formation starts with higher values of δ_r; for example, for $\delta_d = 1.5$, formation first appears for $\delta_r = 4.0$.

In the top right corner of Table 6.1, percentages other than 0 and 100% appear. This goes with a rise in speed of lexical formation. In general, the tendency of the speed of lexical formation is to be high for the lowest values of δ_r with formative success, then decrease to a minimum and to increase again for high values of δ_r.

With respect to the synonymy of the vowel systems that emerged, there are some important observations. First, there is a clear inverse relation between decay rate δ_d and synonymy. Every time you double δ_d, synonymy is effectively halved. Second, the type of relation between reward rate δ_r and synonymy depends on the fact if lexical formation was successful. When it is successful, the relation between δ_r and synonymy is a clear linear one. Every time you double δ_r, synonymy doubles as well. When it is unsuccessful, the influence of δ_r on synonymy is negligible. Instead, δ_d then is the sole predictor of synonymy.

		Reward rate							
		0	0.5	1.0	2.0	4.0	7.0	10.0	15.0
Decay rate	0	0	0	0	0	0	0	0	0
	0.05	0	0	0	13	100	99	74	14
	0.1	0	0	0	34	100	100	100	83
	0.2	0	0	0	57	100	100	100	100
	0.4	0	0	0	76	100	100	100	100
	0.7	0	0	0	0	100	100	100	100
	1.0	0	0	0	0	100	100	100	100
	1.5	0	0	0	0	0	100	100	100

Table 6.1: Percentages of lexical formation for different values of reward and decay rate.

		Reward rate							
		0	0.5	1.0	2.0	4.0	7.0	10.0	15.0
Decay rate	0	51.3	51.3	51.4	51.2	51.2	51.1	51.0	51.1
	0.05	26.5	27.2	27.6	17.7	15.4	24.7	32.3	42.5
	0.1	13.1	13.5	13.7	5.7	7.9	12.7	16.7	22.1
	0.2	6.3	6.5	6.6	2.4	3.9	6.3	8.3	11.1
	0.4	2.8	2.9	2.9	1.3	1.8	3.0	4.0	5.4
	0.7	1.8	1.8	1.8	1.9	1.2	1.5	2.2	3.0
	1.0	1.5	1.5	1.6	1.6	1.0	1.2	1.4	2.1
	1.5	1.1	1.2	1.2	1.2	1.3	1.0	1.1	1.2

Table 6.2: Synonymy of emerged vowel systems for different values of reward and decay rate.

6.4 Discussion

The idea that the formation of a shared lexicon is made possible by the simple processes of learning and forgetting is backed up by simulation evidence in this chapter. In the Introduction chapter, five processes that alter the lexicon were mentioned. The simulation results suggest that if the third process (δ_r) is absent, no lexicon formation occurs. Dito for the fourth process (δ_d). Claims that any other process beyond these four is necessary for lexical formation are hereby invalidated, since formation proved successful without it in this chapter.

Synonym punishment is one of these processes thought to be crucial for forming a shared language. I hope this idea is dropped, also because doing so gives you a richer landscape of possible language systems. In this simulation, we have seen vowel systems with synonymy ranging from 1.0 to 51.4 (and that is just because I capped it off at 50.0). Although these synonymy rates might look inefficient or strange, they offer a more realistic model of reality than the one you get with the addition of synonym punishment. In that case, the synonymy will almost always remain at 1.0, which is not in line with the fact that we do actually have a lot of synonyms and homonyms in the real world.

Quite a few relations emerged from observing the results here. These relations can be caught mathematically in an equation. The maximally supportable and ultimately used number of synonyms is then predicted to be given by

$$s_{max} = \frac{\delta_r}{\lambda R \delta_d}. \tag{6.1}$$

The rationale for having number of referents R in the denominator comes from the fact that agents can only initiate one conversation per time step, while every association is affected by memory decay. Because of this, a higher R results in a lower supportable number of synonyms. λ is a constant that can be determined from the results (it actually lies around 1.0) and accounts for other factors influencing s_{max}. Note that this equation only holds if lexical formation is successful. As you can see, S increases with δ_r and decreases with δ_d.

To test the validity of Equation 6.1, another simulation was set up with fixed values for reward rate, decay rate and number of referents that were chosen beforehand. Table 6.3 shows the expected (based on the equation) and real synonymy found in these simulations. The values are very similar, showing that the equation is indeed valid. In most of the cases, the real s_{max} was a little higher than the expected s_{max}, though. The reason for this could be that agents actually have two opportunities to reward their associations per time step on average: one time when they initiate a conversation and one time when they are invited to converse. So, the maximum expected s_{max} is double the current expected s_{max}. The agents settle for a lower value though, since their communication success is not perfect (thus allowing for a lower s_{max} because of less reward) and creating new associations occurs less frequently. Interestingly, when a synonymy below or around 1 is expected, no shared lexicon is formed by the community in the actual simulation. Apparently, the equation only holds for values of s_{max} above 1.

R	δ_r	δ_d	Expected s_{max}	Real s_{max}
4	5.0	0.1	12.50	13.97
4	5.0	0.5	2.50	2.60
6	5.0	0.1	8.33	9.43
6	5.0	0.5	1.67	1.68
10	5.0	0.1	5.00	6.18
10	5.0	0.5	1.00	1.32
10	5.0	1.0	0.50	1.23

Table 6.3: Testing the maximum synonymy equation: expected and real amount of synonymy for a selection of parameter settings; R is the amount of referents, δ_r is the reward rate and δ_d is the decay rate.

In this chapter, simulation results were reported that support the formation hypothesis of the proposed model of linguistic dynamics. Under parameter settings where both reinforcement and decay were present, chances were high that lexical formation was successful. Under settings in which either of these updating processes was absent, lexical formation was sure to fail.

Chapter 7

Vowel change

The goal of this simulation is to test the hypothesis that alignment causes vowel change. Again, I achieve this by using the DEViL tool, this time by simulating linguistic change. For this chapter, it is important to remember that the parameter δ_a represents alignment.

7.1 Variables

In order to do so, 6 agents speaking the same language were put in one location again. This initial language was the outcome of a lexical formation simulation and can be viewed in Figure 7.1. Other constant parameters are the same as in the previous chapter. δ_r is set to 2.0, δ_d to 0.5. The amount of alignment δ_a is varied, to come to four different simulations with values for δ_a of 0.0, 0.15, 0.25, 0.5, 0.8, 1.5, 2.5 and 5.0 respectively. Each simulation is run 100 times and last for 250.000 rounds of aligning games. Every 1,000 time steps, the situation is analysed, resulting in values for all measures mentioned in the measure session. Of these measures, the lexical distance to the initial language (HLD) is used. Note that to achieve this measure, a referential group of agents that speak the initial language is placed in another location. This group does not interact and their memory is not affected by memory decay. With this setup, the ILD is used as a measure for HLD, since the second location permanently holds the initial language.

7.2 Prediction

Based on the idea that alignment causes vowel change, the prediction is that the HLD increases over time with a speed proportional to δ_a. At the extreme of no alignment, no increase in lexical distance is expected. The lexical distance is expected to converge to a distance similar to the distance of

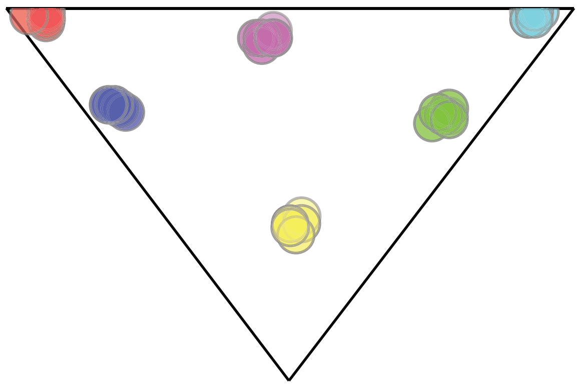

Figure 7.1: The initial vowel space for the community in this simulation. Each bubble represents an association of someone in the community. For each bubble, the following applies: its place is determined by the association's signal place in vowel space and its colour is determined by the association's referent. There is no way to tell whose association a bubble is precisely, but this is not to the point here anyway.

two random languages (LD_{rand}), indicating that it has changed so much that no trace of the initial language can be found. In order to test this, this LD_{rand} is determined by running two simulations for 5,000,000 time steps and taking the average lexical distance from the point that the value lexical distance is settled.

7.3 Results

First of all, the LD_{rand} was determined to be approximately 11.5. There is great variance in lexical distance between two random languages, but still they revolve around this average.

After doing the specified simulations, the results were gathered and passed on to a Matlab-function designed for the purpose of plotting results. This function can be analysed in Appendix 1. The HLD's for the different amounts of alignment are depicted in Figure 7.2. It immediately becomes evident that the distance to the initial language indeed increases when agents align, because all the lines except the one for $\delta_a = 0.0$, have a positive slope. To contrast this, in the case of $\delta_a = 5.0$, the community changed its vowel system to such an extent in "only" 100,000 time steps that it is totally untraceable to the initial system. This is clear simulation evidence for the second hypotheses of the model.

As for the slope of the lines, it does increase monotonically with δ_a, but not exactly proportionally. For every increase in δ_a, the line in the graph gets steeper, meaning that the speed of change is faster (hence the monotonic increase relation). The increase in speed of change between $\delta_a = 0.15$ and $\delta_a = 0.25$ is disproportionately small compared to the increase in speed of change between

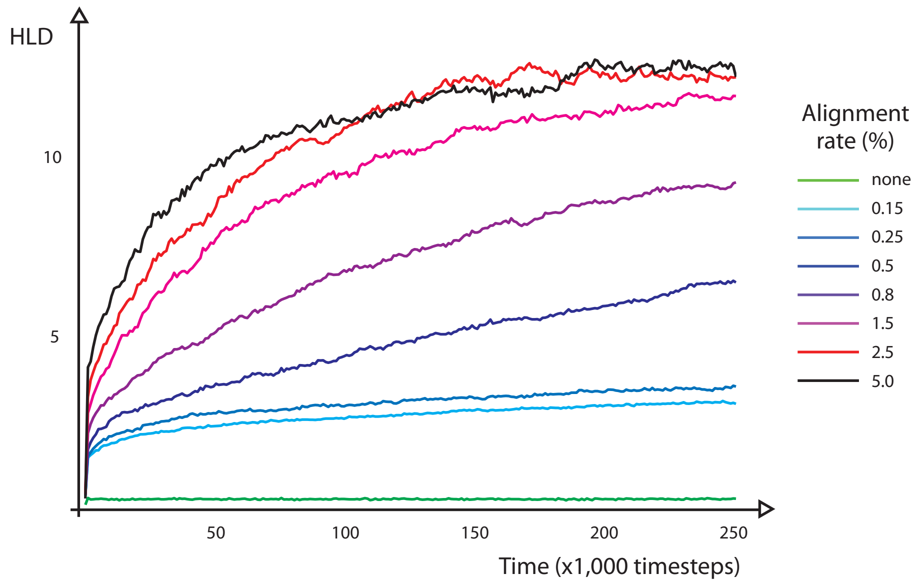

Figure 7.2: Lexical distances to the initial language over time for the different amounts of alignment. HLD stands for historical lexical distance, and measures how much the current language differs from the initial language, its ancestor.

$\delta_a = 0.25$ and $\delta_a = 0.5$, though. Apparently, for this range of δ_a, there are also other forces at play. For higher values, proportionality holds better, indicating that the part of the prediction saying there is a proportional increase relation is true except for low values of δ_a.

As is shown by the lines representing high amounts of alignment, the distance converges to a value that is a little higher than LD_{rand}. This indicates that no trace is to be found of the initial language after a certain amount of time steps and makes the part of the prediction saying that HLD's converge to LD_{rand} true.

7.4 Discussion

Alignment causes languages to change until at some point in time it is not traceable to its predecessor. This point comes faster when there is more aligning. For realistic values of δ_a, this relation is approximately proportional. These values are both not really high and not really low.

Really high values of δ_a are nonsensical and result in either a lot of overshooting or a very unstable lexicon, depending on the implementation of alignment. An example of overshooting is changing your pronunciation of 'love' to 'leave' if someone pronounces it a bit more closed. An example of an unstable lexicon is changing your pronunciation constantly to the last word someone says to you, which would be very strange in any conversation, but especially demeaning in one with several different partners with different dialects. Really low values of δ_a are also in a way nonsensical because the gains of the aligning strategy will then come long after the conversation has ended. This renders employing the strategy to have a low utility for an agent, so in that case it can just as well not align. Even if aligning happens unintentionally, one could argue that for it to have evolved, it should have at least some function.

In this chapter, simulation results were reported that support the change hypothesis of the proposed model of linguistic dynamics. Under parameter settings where alignment was present, change occurred with a speed roughly proportional to alignment rate. Under settings in which this updating process was absent, the language was bound to remain static.

Chapter 8

Vowel divergence

The goal of this simulation is to test the hypothesis that geographical constraints cause vowel divergence. What is known from the vowel change simulation is that over time, alignment to signal variation introduced by production noise causes a language to have an increasingly higher distance to its initial state, until it revolves around the average distance between random languages. In this chapter, the DEViL is exploited again to simulate agents in two neighbouring communities changing their vowel systems simultaneously, while traveling to and fro.

8.1 Variables

Therefore, in this simulation, a second location is added. A community of agents identical to the one in the first location is placed in this location. These communities are in turn identical to the ones used in the previous experiment. They also utilise the vowel space of Figure 7.1. The constant parameters are identical to the previous chapter too. δ_a is set to 1.0. The amount of traveling τ between the two locations is manipulated to the values of 0, 10, 20, 50, 100, 200, 500 and 1000, to see how this affects the resulting system behaviour. Remember that τ is implemented as the probability (in parts per million) of an agent deciding to travel each time step. Just as a quick example, under the setting $\tau = 1000$, each agent would have a 1000 ppm probability of traveling each time step, which amounts to a 0.1% chance per agent per time step. In a simulation with 12 agents such as the current one, this results in an approximate grand total of one travel per 88 time steps. Analogously, under the setting $\tau = 10$, there is one travel per 8800 time steps on average. All simulation settings are run hundredfold for 250,000 time steps.

As a measure, the interlocal lexical distance (ILD) is used. In short, this measure is a number that becomes higher when the local lexicons becomes more different. The ILD is zero at the start of the simulation, since the communities of the two locations speak exactly the same language.

8.2 Prediction

What is predicted is first, that the interlocal distance increases over time. This follows from the results of last simulation and the fact that the direction of change is not predetermined and is therefore interlocally independent. Second, it is predicted that the speed of increase decreases with an increase in traveling (so an inverse relation between τ and increase speed). When there is no traveling, the amount of vowel divergence is predicted to be at its maximum, the amount gradually decreasing when start people traveling more. At the other extreme, when every agent travels all the time, no vowel divergence is expected to arise at all, since in that case the two locations do not act as dividers between the two communities. Instead, there is in that case expected to emerge a single population consisting of the two combined communities that are located at the same time in both locations.

8.3 Results

After doing the specified simulations, the results were gathered and passed on to the Matlab-function mentioned before and available for viewing in Appendix 1. Figure 8.1 depicts the ILD's over time for the different amounts of traveling. First of all, the ILD does indeed increase over time, since again the slope of the lines are all positive. This is important, since this is simulation evidence for the third hypothesis of the proposed model of linguistic dynamics. This completes the body of evidence for the model, with every hypothesis now being supported by simulation evidence.

Without traveling, the ILD rises quickly to converge to LD_{rand}. The same behaviour is seen for values of τ up to and including 20. For $\tau > 50$, a different picture is observed: the ILD first increases but then settles for a value lower than LD_{rand}. In case of $\tau = 50$, this ILD settling value seems to lay around 7. A simulation with more time steps needs to be run to justify whether this particular parameter setting settles or very slowly increases. A travel rate of 100 results in an ILD of just below 5, while even higher travel rates settle for an ILD of 3. Somehow, values of ILD below 3 do not occur from the start. This should be attributed to the implementation of the ILD measure, which rates very similar vowel spaces (for example only with different association strengths) to be already significantly more different than identical ones. The figure in the previous chapter showed this phenomenon as well.

The simulations results are thus in accordance the predictions in several ways. The ILD indeed

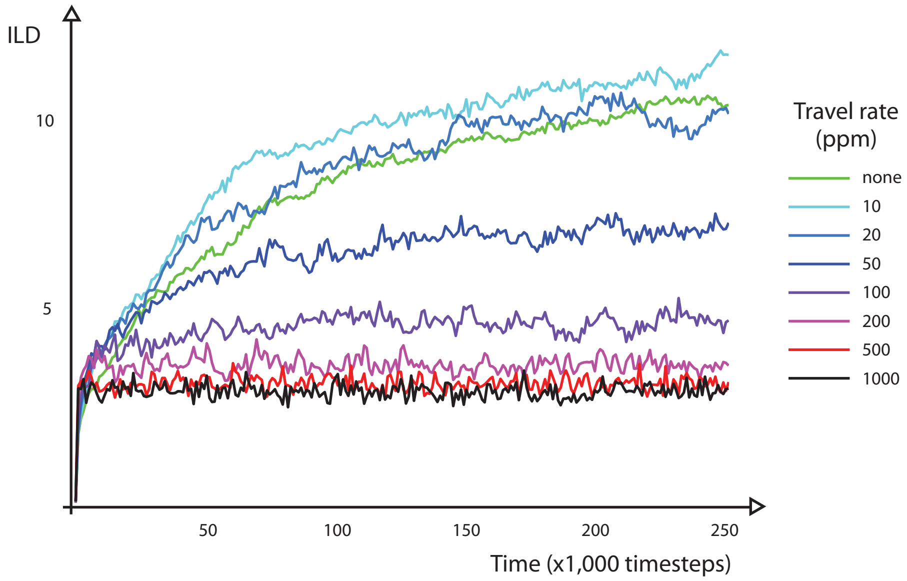

Figure 8.1: ILD over time for the different amounts of traveling. ILD stands for interlocal lexical distance, and it measures how much the two local languages differ.

increases over time. Furthermore, for $\tau > 50$, the speed of increase can indeed be said to decrease, since the slopes of the lines representing these travel rates get increasingly moderate. Perhaps more interestingly, the simulation results indicate that for these travel rates, the ILD settles for an increasingly lower value, instead of converging to LD_{rand} as predicted. Note that this is not in conflict with the model or any of its hypotheses: rather it should be viewed as a refinement of its predicted behaviour. How so will be discussed in detail in the following discussion and in the General discussion chapter.

8.4 Discussion

The languages of two neighbouring locations tend to diverge. If no traveling occurs, the languages quickly diverge to LD_{rand}. Traveling has a negative effect on this divergence: the more traveling occurs, the less the languages diverge. Interestingly, for "medium" amounts of traveling, languages settle for a lower ILD than LD_{rand}. Why does this happen? I explain. Figure 8.2 gives a visual impression of the scale of linguistic distance. On it, ILD_{int} marks the turning point between intelligibility and unintelligibility. When an agent travels to the other location, where the vowels used to signal the referents are different, two scenario's can occur, depending on the mutual intelligibility between the agent's language and the new location's language. In other words, if $ILD < ILD_{int}$, a different scenario will ensue than if $ILD > ILD_{int}$.

If the languages are mutually intelligible, interactions between the newcomer and the resident

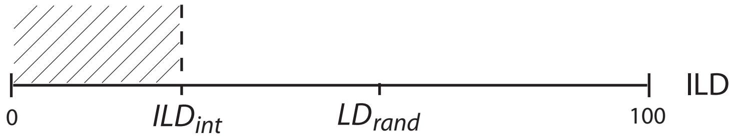

Figure 8.2: Abstract scale of ILD with two important points indicated: ILD_{int} and LD_{rand}. Languages differing less than ILD_{int} are mutually intelligible and are therefore prone to aligning, as indicated by the shaded area. LD_{rand} is the average distance one finds when comparing two randomly picked language. This constant has proven to lie around 11.5 for the settings used in this chapter.

community will be successful and thereby prone to alignment. The traveler's language will change towards the local language via the alignment process during these interactions. If the languages are mutually unintelligible, no alignment will occur since interactions are unsuccessful. Instead, the agent's old language falls prey to punishment of disuse and thus gradually disappears. At the same time, it simply takes over the language of the local community through the process of lexical formation. In summary, when you travel there is a distinction between aligning to and learning a new language, depending on how different the language is to yours.

On the scale in Figure 8.2, the area in which alignment occurs is shaded and the area in which learning occurs is not. If the ILD is within this area, the traveling agent will use alignment to change its manner of pronunciation to the one used in that location. A side effect of this alignment is that it works both ways: every time the agent is the speaker and an interaction is successful, the hearer, which is a local inhabitant, will align its pronunciation towards that of the newcomer. The local language is thus affected by the traveler, effectively reducing the interlocal lexical distance until the traveler's lexicon is similar to the local one. If the interlocal lexical distance is not within this area when an agent travels, the agent will use learning to change its manner of pronunciation. This strategy does not have the side effect that alignment has, meaning that in this case the ILD is not affected by the traveler. So, each time an agent travels when ILD is below ILD_{int}, the ILD decreases.

This means there are really two processes going on in this simulation: first there is the general increase in lexical distance over time caused by vowel change, in turn caused by alignment to signal variation and second there is a potential decrease in lexical distance each time an agent switches location. Let us use this knowledge to explain the behaviour we observe in the simulation. The initial ILD is zero: the two locations' inhabitants speak the same language. Right at the moment the simulation starts, the first process sets in and the ILD starts to increase, slowly but surely diverging the languages of the two locations. In this stadium, when an agent travels, the second process will cause the ILD to drop again. This continues until the system reaches a particular steady state.

These steady states are one of the subjects in Chapter 10.

In this chapter, simulation results were reported that support the divergence hypothesis of the proposed model of linguistic dynamics. Under parameter settings where geographical constraints were present, languages diverged with a speed roughly proportional to the severity of the constraints. This Under settings in which this updating process was absent, the language was bound to remain static.

Chapter 9

Further analysis

There are several interesting questions to ask about the behaviour of the system of agents, the aligning game and the map of locations, because this system is complex and adaptive. In this chapter, two of these questions are covered, namely:

Question one What parameters affect formation success?

Question two What are the effects of noise?

These questions are answered in considerable detail, even making use of additional ad hoc simulations.

9.1 Parameters affecting formation success

Chapter 6 showed us that simulations in which all five updating processes were present (including reinforcement of communicative success and punishment of disuse) resulted in formation success. Yet, not every simulation containing the five processes will lead to successful formation. The whole set of parameters, including A, R, δ_d, δ_r, Ψ_{ac}, and ω_0, is responsible for laying down constraints for the five processes that need to be met in order for the system to come out successfully. Some parameter configurations allow for easy lexical formation, while others render it very difficult or even impossible. In Table 6.1 on page 32 for example, we see that low rates of reward are not sufficient for formation success. In the case of the settings chosen for this simulation, values of reward rate below 2.0 are apparently insufficiently low. In this section, it is analysed how the aforementioned parameters affect the successfulness of lexical formation.

Although the amount of parameters make this particular system complex and hard to formally analyse, it is possible to indicate how each parameter affects formation success. During lexical formation, there can be said to exist four phases, namely (1) the creation phase, (2) the propagation phase, (3) the competition phase, and (4) the maintenance phase. Imagine the community of agents at the start of a simulation, when all agent have empty lexicons. This is when the creation phase begins. After the first game, the speaker creates an association and the hearer takes it over. In subsequent games, the odds of a speaker creating an association slowly decrease because its lexicon is bound to fill up with associations. At a certain point, the creation phase can be said to be over, and the speakers do not create any new associations. Then the created associations are propagated through interactions to the whole community. When every individual has every association, these associations have to make out which one will be used in the ultimate shared language in the competition phase. When this phase is over, a shared language has emerged, and it is simply maintained in the maintenance phase. There are exceptions to this four-phase-system: some systems enable a very fast transition from the first to the fourth phase (e.g., those with high reward rates), other systems prohibit progress to the fourth phase (e.g., those with high decay rates), and again other systems skip the third phase altogether (e.g., those which allow a highly synonymous lexicon to be maintainable).

Parameters have different effects in different phases. Furthermore, the outcome of one phase can influence what will happen in the next phase. For example, when the synonymy after the propagation phase is very high, it might be harder for the system to develop a successful, shared lexicon. For this reason, it is important to know what factors play in early phases, before answering how parameters affect formation success. That is why we first regard the number of associations that are created during the creation and propagation phase. More concretely, what is the expected synonymy after propagation has ended? This synonymy appears to depend on the amount of agents A.

If $A = 2$, synonymy will always be 1, because for each referent one agent creates a vowel and the other takes it over, thereby preventing the latter from making its own competitor. For $A > 3$, things start to get out of hand. Table 9.1 provides a way to compute the expected synonymy for higher number of agents, in which P is the number of agents having a signal for a referent. The computation focuses on one referent, because synonymy is identical across referents. For each number of agents, you can create a tree of possibilities starting with $P = 0$ and $s = 0$ using the probabilities and effects given in the table. An example tree is shown in Figure 9.1. Doing this, you find that each time you increase A by 1, expected synonymy is increased by 0.5, so $s = \frac{1}{2}A$.

In the second phase, these candidates will be propagated to each agent, so that all agents have $\frac{1}{2}A$ vowels for each referent, all with the initial associative strength $\omega_\alpha = \omega_0$. For a shared, successful lexicon to emerge, one of these candidate associations needs to strengthen to ω_{max} at the expense of its fellow candidates. This battle is fought in the competition phase. We now come to our initial

Event	p(Event)	Effect
strengthen	$\frac{P^2-P}{A^2-A}$	—
expand	$\frac{-P^2+AP}{A^2-A}$	P+1
new to old	$\frac{-P^2+AP}{A^2-A}$	P+1, s+1
new to new	$1-\frac{-P^2-2AP+P}{A^2-A}$	P+2, s+1

Table 9.1: Four events that can occur after an aligning game in the creation phase. When an agent that already has a word for a referent is next to speak and chooses another agent that already knows the referent, a 'strengthen' event occurs; when an agent that already knows the referent chooses an agent that does not, an 'expand' event occurs; when an agent that does not know the referent chooses one that does, a 'new to old' event occurs; when an agent that does not know the referent chooses one that does not either, a 'new to new' event occurs. A is the number of agents, P the number of agents that have a signal associated with the specific referent, s is the synonymy of the specific referent. For each event, the probability of it occurring is given, as are the effects on P and s after the occurrence.

question: what is the chance of this formation of a shared, successful lexicon succeeding?

To approximate this chance, we first define associative flux as the expected change of associative strength in a candidate association x after one time step. The associative flux is given by

$$\Phi_F = r_x\delta_r - \delta_d, \tag{9.1}$$

in which r_x is the expected number of successful aligning games involving the candidate association per time step. r_x is dependent on the probability of it being chosen as topic $p(c_x)$ and the probability of it being guessed right $p(g_x)$.

In noiseless environments (here we look at systems without noise; the effects of noise are discussed in the next section), the topic is always guessed right—these agents are artificially intelligent, remember—making $p(g_x) = 1$. The probability of a candidate being chosen is given by

$$p(c_x) = \frac{\omega_x}{R\sum_{\alpha\subset\aleph_x}\omega_\alpha}. \tag{9.2}$$

In this equation, ω_c is the candidate's associative strength, $\aleph_x$ is the set of associations that have the same referent as the candidate (itself and its competition, so to say) and ω_α the associative strength of association α in that set. The equation follows from the fact that a referent is chosen randomly, making the probability of the candidates referent being picked $\frac{1}{R}$ and the fact that a vowel is chosen based on its relative strength.

Each agent initiates an aligning game each time step, so the number of games per time step is A.

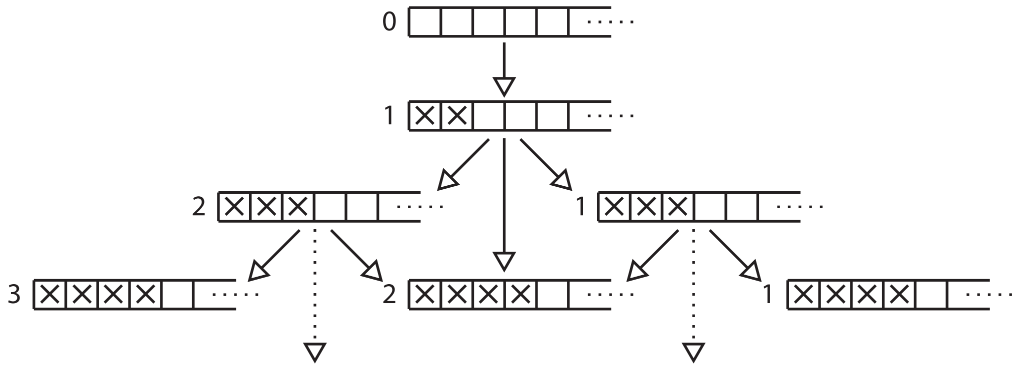

Figure 9.1: Tree of possibilities for the creation phase. The row represents the community of agents, each space representing an agent. P can be read as the number of crosses in the row, and s is the number to the left of the row. An arrow downwards indicates a 'new to new' event, in which a new signal is introduced and instantly propagated to an agent that does not know the referent; an arrow to the left indicates a 'new to old' event, in which a new signal is introduced; an arrow to the right indicates an 'expand' event, in which an old signal is propagated. If A is known, the probabilities of each event can be computed. After a nested multiplication, s can be computed. Note that the 'strengthen' event has no arrow, since the system is not affected by it.

This in turn gives us an expected number of successful games of

$$r_x = Ap(c_x) = \frac{A\omega_x}{R\sum_{\alpha \subset \aleph_x} \omega_\alpha} = \frac{2}{R}. \tag{9.3}$$

The last step can be made, because each association has the same associative strength and $s = \frac{1}{2}A$.

If one now fills in the associative flux equation for $R = 6$, as was used in Chapter 6, we get a flux of $\frac{1}{3}\delta_r - \delta_d$. So, for a system to be in constant flux in these circumstances, δ_r should be chosen three times as high as δ_d. For higher reward rates, the flux is positive, providing room for a candidate to rise to ω_{max}. For higher decay rates, this becomes impossible, as can be seen in Table 6.1 on page 32, in which formation never succeeds for values of δ_r less than 3 times δ_d. What now happens when one candidate is reinforced once, making its association strength $\omega_0 + \delta_r$ and its competitors still ω_0? More concretely, what is the increase in flux after communicative success? Lets further assume that there are six agents, as in Chapter 6, and that $\delta_r = 3$ and $\delta_d = 1$ (creating a constant flux).

In case of $\omega_0 = 10$, the flux of the lucky candidate becomes 2/11, whereas the flux of its competitors drops to -1/11. In case of $\omega_0 = 1$ however, the flux of the candidate becomes a staggering 1, whereas the flux of its competitors drops to -1/2. If ω_0 gets lower (relative to δ_r), the increase in flux becomes larger. Larger flux increases after communicative success mean that agents need to complete less successful aligning games in order to develop a shared language. We can conclude from these calculations that high ω_0 values thwart successful lexical formation. In retrospect, $\omega_0 = 10$ might have been a high value to use by default. Still, from $\delta_d = 2$ onwards, lexical formation already starts to succeed, indicating that a small edge in flux of strength can be decisive in the long run.

Increasing the number of agents to $A = 100$ leads to a flux increase after communicative success of $\Delta\Phi_F = 147/503$ for $\omega_0 = 10$ and $\Delta\Phi_F = 200/53$ for $\omega_0 = 1$, both of which are a higher increase than for six agents. There are 100/2=50 competitors after the propagation phase though, which means a candidate needs to beat much more other associations to win the competition. Overall, the effect of increasing A is thus probably negative, especially when noise is introduced.

Increasing the number of referents to $R = 100$ leads to a flux decrease of $\Delta\Phi_F = -511/550$ for $\omega_0 = 10$ and $\Delta\Phi_F = -22/25$ for $\omega_0 = 1$. This is in agreement with Equation 6.1 on page 32, which dictates that increasing R lead to a decreasing number of sustainable synonyms. For $R = 100$, this number is presumably far below 1.0. To answer the first question of this chapter: all things equal, increasing δ_r raises formative success, whereas increasing R, δ_d or ω_0 lowers it. Increasing A can be said to have multiple effects: it increases the amount of competitors after the propagation phase, which makes it harder for a shared language to develop, but it does give a slight flux increase, which helps formative success.

9.2 Noise and its effects

In this section, the effects of acoustic noise on the system behaviour are discussed. When there is no variation in the produced signals, an entirely different behaviour on an individual and communal level is observed than when there is a lot of production noise. Also, the resulting vowel systems are of a different nature, depending on the amount of noise in the simulation world. Of course, the direct effect of added acoustic noise is variation in produced vowels. This is variation in the external language (the E-language). Agents intend to produce a prototype vowel, but they produce one that is a variation of this prototype (see Figure 5.2 on page 28). At least five behavioural differences result from this.

The first effect of noise relates to the taking over process. When a hearer takes over an association between vowel and referent, the only thing it can rely on is the vowel that has actually been uttered. Its best guess is to use that vowel in its new association. If there is no noise, this vowel is identical to the intended prototype, ensuring that the vowel taken over is also the exact same prototype. When there is noise however, the hearer is likely to take over a variation of the prototype, resulting in variation in the internal lexicons of the community. So even when there is no alignment, the process of taking over associations is a cause of variation in the I-language.

The second effect of noise is similar to the first Analogously, aligning can only rely on the uttered signal. So, agents do not align to another agent's prototype, but instead to their utterances, which, when noise is present, vary in their place in vowel space. Imagine two agents with identical lexicons. With no noise, one agent will always pronounce exactly the intended signal. E-language and thereby I-language thus remain constant in this scenario. With noise however, the speaker produces a variation of the prototype and the hearer aligns to this, resulting in a change of its I-language. Now that the hearer's signal has shifted in vowel space, its subsequent utterances will center around this new prototype. There is not just one altered replication of the utterance: there is what Croft (2000) calls a differential replication. A differential replication is a shift in the frequencies of variants of a vowel. This is the case here, since the agent who adjusted its prototype will now produce a shifted distribution of utterances for that vowel, based on the new prototype and acoustic noise. Acoustic noise can thus be said to provide markers for alignment to move towards, leading to an continuous interplay between shifting I-languages that itself change the expected E-language, again shifting I-languages, and so on.

The third effect of noise is different. Whereas the last two were ways in which the system becomes less static, this difference pertains to conversation successfulness. Imagine again the two agents from last paragraph. Now, when the agents play an aligning game without noise, the hearer will always guess the topic right, since the speaker's prototype is identical to the spoken utterance and to the hearer's prototype. When searching for low lexical distance, identity is always closer than a positive distance, making communication successfulness a certain 100%. When they play a game with noise, the spoken utterance is not the same as the hearer's or speaker's prototype. Therefore, it is suddenly not certain if the best choice for the hearer is also the right choice. It could be that the utterance is closer to another vowel than to the intended one, making that the best guess for the hearer. A communication successfulness below 100% is inevitable because of this. Also, the higher the amount of noise, the higher the odds of failure. For really high amounts of noise, communication success even falls to the level of chance.

The fourth effect of noise follows from the third. When successfulness is below 100%, the assumption that $p(g_x) = 1$ does not hold. So in that case, Equation 9.3 is insufficient in predicting the number of successful aligning games per time step. Instead, the actual flux will always turn out to be lower than what the equation proposed earlier predicts, because $p(g_x) < 1$. A lower flux leads to two things: vowel systems become harder to form and costlier to maintain. As mentioned in the previous section, for an association to rise to the top, it needs to have a positive flux; there must be more associative strength going in than out. Failing games systematically reduce the strength going in and with it the chance of success for an association. So, ceteris paribus, increasing noise to a simulation leads to a more negatively inclined flux equation for each association and thus potentially to smaller odds of lexical formation. Equation 6.1 was an early effort of formalising how many synonyms an agent can maintain, given the model parameters. In hindsight, I assumed perfect game successfulness. Actually, δ_r needs to be multiplied with $p(g_x)$ in this equation, so that the maximally sustainable synonymy decreases with increasing failure and noise rates. The two equations mentioned in this paragraph go together intimately.

The fifth and final effect of noise is that it pushes communities towards an optimal use of the signal space, in this case the vowel space. Optimal vowel systems become attractors (de Boer, 2001) in systems with a sufficient amount of noise. An optimal vowel system makes use of the vowel space by spreading out the individual vowels as much as possible. The most optimal vowel system with 6 vowels consists of the three quantal vowels /a i u/, since they are in the extremes of the vowel space, and /ɨ ɛ ɔ/, which lay between pairs of quantal vowels. In Figure 9.3, the most optimal system can be appreciated in vowel space type D. Any other type of vowel space will have at least one pair of vowels for which the distance in vowel space between them is lower than the smallest distance in the optimal system. Given a sufficient amount of noise, each distance reduction has more failing aligning games as a consequence. Failing games in turn lead to a lower flux. The precise steps as to why a slightly suboptimal shifts to a the optimal one in most cases are quite complex and are reviewed in some detail in the section called Vowels as magnets and atoms. What it comes down to is that on the level as "low" as that of aligning games and utterances, associations that work relatively badly are punished on the long run. Even with only slightly lower success odds, an association will just decay more and therefore be chosen less often, making it decay even more until it eventually dies out.

A quick summary of the effects of noise. Besides the intended effect of adding variation in the E-language, the effects are:

1. variation in the I-language because of taking over,
2. variation in the I-language because of alignment,
3. communicative success below 100%,
4. low chance of formation/systems harder to maintain, and
5. vowel space is used optimally.

The model proposed in the Chapter 1 can be expanded upon to include these effects. The result of this expansion is shown in Figure 9.2.

To verify what has been said, I run a post hoc simulation in which all parameters are fixed except for acoustic noise, which varies from 0.0 to 30.0 with steps of 2.0. δ_r was set to 10.0, δ_d to 1.5 and

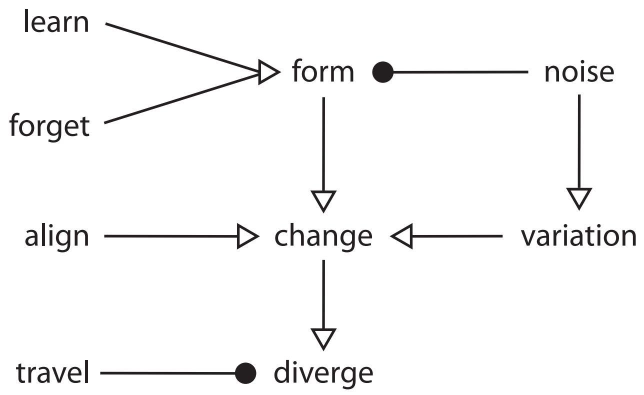

Figure 9.2: Abstract diagram of the model of linguistic dynamics, including two of the effects of acoustic noise.

δ_a to 0.1. The other settings were like the lexical formation experiment, so there were 6 agents without an initial language in one location. Based on the five effects of noise, it is expected that for low amounts of noise, there is low local variation LV, high communicative success c, low formation success, low synonymy s and suboptimal use of the vowel space, whereas for high amounts of noise, there is the opposite for each of these measures.

Table 9.2 displays the results of this simulation. In this table, I added a column with information about the quality of the vowel system. I observed six general types of systems that emerged in the different simulations. These were systems with (A) its vowels randomly positioned, with at least one vowel pair having a distance closer than 15% of the vowel space, (B) its vowels randomly positioned, but with all its vowel pairs having at least a distance to each other of 15% of the vowel space, (C) four high vowels /i ɪ ʊ u/ and two middle vowels /ɛ ɔ/, (D) three high vowels /i ɨ u/, two middle vowels /ɛ ɔ/ and a low vowel /a/, (E) only the three quantal vowels /i a u/ and (F) no vowels at all. Type D is the most optimal system. It is in accordance with Crothers' hierarchy, indicating that it occurs very often in the real world, whereas type C is too much inclined to high vowels (/ɪ ʊ/ should only be introduced after /a ɛ ɔ e o/). Example visualisations of all six types of vowel systems can be seen in Figure 9.3.

In table 9.2, it can be seen that low levels of noise lead to random vowel systems of type A moving on B. At the level $\Psi_{ac} = 6.0$, type C and D systems are introduced. For $\Psi_{ac} > 10.0$, we only see type D systems, which turn into type E systems at $\Psi_{ac} = 18.0$. With a noise level in the high 20's, we see only type F systems. The table also shows that s drops from 100.0% to 22.9% (100-failure). Formation success only occurs for $\Psi_{ac} < 18.0$ and occurs faster for lower noise levels. Also, local variance goes up (from 0.6 to 31.9), synonymy increases (from 1 to 1.5) and strength decreases (from 90.2 to 8.6).

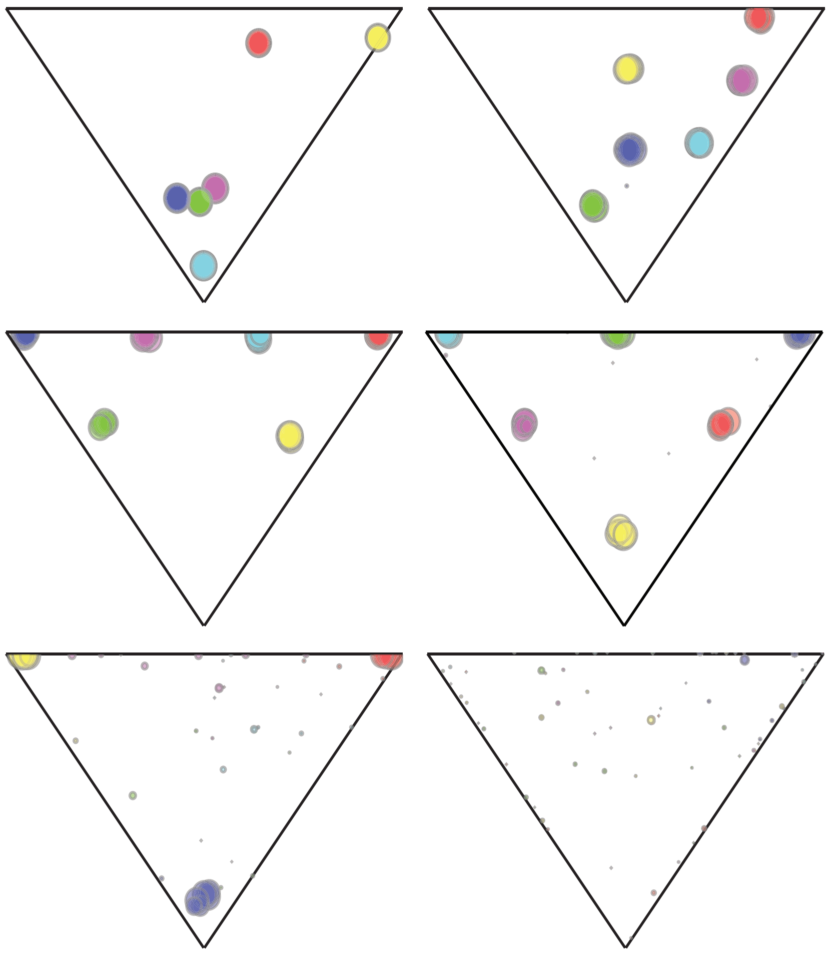

Figure 9.3: Visual examples of the six vowel system types that arise for different levels of acoustic noise. From left to right, from top to bottom, the examples are in alphabetical order of type.

		Form. speed	Fails (%)	Lex. quality
Noise (σ^2)	0	110	0.0	A(8) B(2)
	2.0	150	0.1	A(8) B(2)
	4.0	240	0.6	B
	6.0	450	1.9	B(3) C(3) D(4)
	8.0	700	4.2	C(1) D(9)
	10.0	2,230	9.4	C(2) D(8)
	12.0	3,360	14.0	D
	14.0	10,010	21.3	D
	16.0	>100,000	27.2	D
	18.0	x	49.0	E
	20.0	x	53.3	E
	22.0	x	56.8	E
	24.0	x	61.1	E
	26.0	x	68.0	E(8) F(2)
	28.0	x	74.2	E(2) F(8)
	30.0	x	77.1	F

Table 9.2: Influence of noise on lexical formation and quality of shared lexicon

All these results back up the analyses made in this section. That noise leads to optimal use of the vowel space comes forward in the fact that type D systems only arise in the lexical formation fostering simulations in the higher range of acoustic noise level. Type E systems, which are optimal three-vowel systems, persist for the levels of noise just above the ones allowing formation, indicating that agents try to make the most of the space even when they cannot form a complete shared language. The fact that we only see type F systems for the highest noise levels supports the idea that noise lowers the chance of formation. The dropping successfulness and the enormous increase in local variance are also conform expectations. The only thing that contrasts with the predictions is the increase in synonymy. Perhaps the fact that failing games lead to agents taking over associations, which in turn gives synonymy a boost, was overlooked. These newly created associations remain at low strength, but do contribute to a higher synonymy. Either way, synonymy did not vary much under these settings (from 1.0 to 1.5); the fourth effect is presumably more relevant for higher reward-decay ratios.

Chapter 10

General discussion

In this study, the goals were to create a model of linguistic dynamics and a simulation tool to verify this model's validity. In Chapter 4, the proposed model of linguistic dynamics was described. The three important hypotheses that were derived from the model were that learning and forgetting drives lexical formation, alignment drives linguistic change, and geographical constraints drive linguistic divergence. In Chapter 5, a description was given of the tool that was designed and implemented to test the model. The agents in this tool's simulations played aligning games, which are a new type of language game, based on the proposed model. This game permits five updating processes to alter agents' lexicons: creation, taking over, reinforcement and forgetting of an association and alignment (see Chapter 4 for a more detailed description). In the three chapters that followed, simulations were run to test each one of the three hypotheses.

What do the results of these simulations say about the model and about linguistic dynamics in general? The tendency seen throughout the simulations is that they confirm the model, but that they also provide new insights, which are analysed in Chapter 9. These insights can be used to update the model and to come up with new paths in tackling the challenge of modelling linguistic dynamics.

The formation hypothesis stated that lexical formation is driven by reinforcement of communicative success and punishment of disuse. In the simulation of Chapter 6, lexical formation proved to be possible under parameter settings involving positive reward and decay rates. When all five lexicon altering processes were in place, lexical formation had a high chance of succeeding. Other processes, like various types of punishment, are thus not necessary. This contrasts for example with the idea that synonym punishment is a vital process in the formation of a shared lexicon, proposed by Steels (1999). Instead of being vital, these processes should be seen as playing additional roles at best. In Chapter 9, it appeared the criteria for formation success were more complex than simply the presence of reinforcement and memory decay. First, the amount of reward must be high enough compared to the amount of memory decay for a lexicon to be maintainable and lexical formation to be successful. Additionally, high numbers of referents and high initial associative strength proved to negatively influence formation success. Finally, a high number of agents produced a high number of synonyms in the early phase, which made it hard for a winning association to arise in a later phase, thereby making lexical formation difficult.

In the simulation of Chapter 7, alignment proved to be a driving factor in linguistic change. This is in alignment with what the alignment hypothesis predicts, thereby validating this hypothesis. It was seen that a language starts becoming more distant from its predecessor language, until it eventually becomes so different that no trace of the predecessor language is left. This point was called LD_{rand}, indicating the average lexical distance of two independent, randomly picked languages. The time it takes for a language to differ LD_{rand} from its predecessor depends on the amount of aligning: the more aligning, the faster the change, and the quicker no trace is left. This does not imply that alignment is the only reason that languages change. In fact, it would be interesting to model alternative mechanism, such as the ones described in Chapter 2 computationally to find out whether they too hold up under the stress of explicit simulation. This could lead to better insights into which mechanisms play which role in the process of linguistic change.

In the simulation of Chapter 8, geographical constraints were shown to be a driving force in linguistic divergence, thereby proving the divergence hypothesis. We saw that agents in two locations with the same initial language move in different directions in the vowel space, effectively moving the languages apart. This linguistic divergence is dampened or inhibited by travel between the two locations. In extremum, when every agent travels frequently, there is no distinction left between the two locations, so that they can be seen as one. The initial language then also moves in the same direction for both locations, the lexical distance between local languages remaining small. What this means is that when part of a community with a shared language moves to a new location and rarely travel to and fro after this move, the communal languages will uncontrollably start to differentiate. This is what happened to the English language after the colonisation of the New World, and what caused people to sing about saying to-mah-to and to-may-to.

Acoustic noise was shown to have several effects, five of which were recognised and listed. These were (1) linguistic change induced by the updating process in which agents take over associations, (2) linguistic change induced by alignment, (3) a decrease in communicative success, (4) vowel systems having higher maintenance costs, and (5) the vowel space being used optimally. These effects all sound like they exist in reality as well. When learning new words or adjusting your speech, you can not make use of the internal linguistic representations of the one you learn from or talk with. Instead you rely on explicit utterances. If these vary, your learning and aligning varies depending on this variation in the E-language. The result of this is that your I-language is not identical to that of your teacher or interlocutor, the result of which is that your E-language will be different in the future: linguistic change.

It is more difficult to relate effects 3, 4, and 5 with reality, without proceeding into talk about alternate worlds without any noise. One could meditate that communities living in conditions with exceptionally high environmental noise, such as a construction site—please bear with me—, have a higher need for an optimally used vowel space in order to have at least some communicative success. What is actually known, though, is that there is acoustic variation in the real world, that communication is not always successful, that one can not maintain an infinitily large vowel system, and that vowel systems tend to use the vowel space optimally. This thesis' simulation results link these observations together by showing that for certain degrees of acoustic noise, this realistic behaviour emerges.

In Chapter 9, a division of lexical formation in four phases was given. These four phases were the creation, the propagation, the competition and the maintenance phase. Although the bulk of today's languages were formed a long time ago and therefore reside in the maintenance phase, the other three phases are also still active today and in the near past in for example Pidgin and Creole languages, and to a degree in the creation of sign languages, individual acquisition and linguistic invention.

Communal languages, and with that the system state, are different in each of these phases. In the creation and propagation phases, synonymy is high, associative strength is low, communicative success is usually low and local variance is high. After a successful competition phase, synonymy is low, associative strength is high, communicative success is high and local variance low. These two states are roughly represented by type A and type D systems in Figure 9.3.

Some of these phase related states are steady states, while others are transitional. When a system steamrolls through the first three phases to the maintenance phase, the states it passes on its way are of course transitional. The final state, as described in the previous paragraph, can be called a steady state. Another steady state occurs when the decay rate is too high for the reward rate. In this case, the system remains in a type A like state.

In extreme cases, other states are possible. For example, with enormous levels of noise, reward and alignment (a situation perhaps resembling a very loud environment where life and death decisions

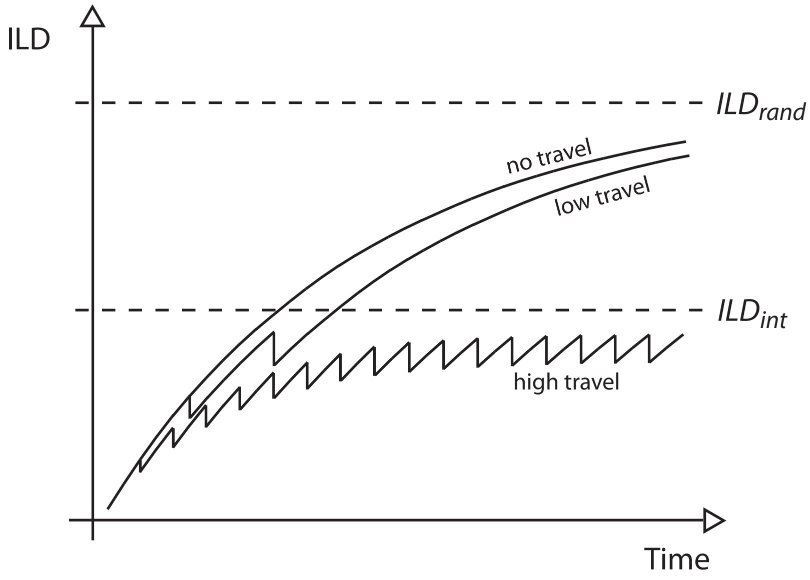

Figure 10.1: Given two groups with the same initial language, one can predict whether their languages remain identical or become dialects or different languages from the amount of traveling between the groups. When no traveling occurs, the languages diverge to ILD_{rand}, the mean ILD for two languages picked at random; when little traveling occurs, the language eventually diverge beyond ILD_{int}, the threshold of intelligibility, to become different languages; when enough traveling occurs, the languages fail to break through ILD_{int} and thus remain at the level of dialects. If even more traveling occurs, the languages will remain the same.

have to be made based on interaction), agents seem to come up with a fast changing shared lexicon using the three quantal vowels. In this scenario, the quantals can be seen as homophones describing more than one object. Since making this three way distinction is better than no distinction at all, this system is probably the optimal one given the circumstances.

The states that languages of two locations in relation to each other can be in are categorisable as well. Three possible steady (end)states are theoretically possible and observed in the simulations in Chapter 8. The first, $S1$, is trivial: the languages of the two locations are equal for all time and can just as easily be seen as one language. The second, $S2$, is more complex: the ILD settles at a value below the turning point of intelligibility ILD_{int} (see Figure 8.2 on page 39) and the two languages can be seen roughly as dialects. The third, $S3$, is as follows: the ILD continues to rise until it revolves around the LD_{rand}; the two languages can then roughly be seen as two different languages.

In the simulation, which specific end state is reached critically depends on the parameter that regulates the amount of traveling. The nature of this dependence is depicted in Figure 10.1. If no traveling occurs, the third end state is reached. From a certain amount of traveling, we arrive in the second end state. For this to happen, the amount of traveling should be high enough to compensate

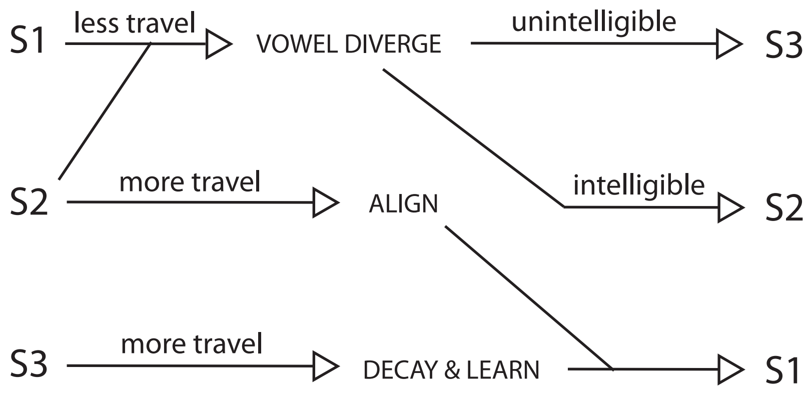

Figure 10.2: States of two local languages and their transitions.

the amount of increase in ILD caused by the first process between each travel. If too little agents travel, the ILD crosses ILD_{int}, after which alignment has no effect and the ILD continues to rise as proposed for the third end state. The more agents travel, the lower the ILD at which the simulation settles, eventually getting so low that the first end state can be said to be reached. If all agents travel all the time, the geographical constraint is so low that the two locations just as easily can be seen as one, since all agents interact with each other in this case.

In reality, geography and demographics are dynamic. Natural boundaries, geographical constraints, and social/political borders come and go. Figure 10.2 shows how the model predicts the state of two local languages to change in reaction to a change in travel rate (which in turn can be caused by an altered geographical situation). In this figure, it can be seen that when a system is in state $S3$ and the amount of travelling increases, the system will undergo a state transition under the influence of decay and reinforcement. The system then ends up in state $S1$. These predictions agree well with observations made in the real world; they help explain many scenario's in historical linguistics such as the difference between to-mah-to and to-may-to.

An interesting analogy that also pertains to this study is the view of words as magnets (de Boer, 2001, p. 14). One could imagine that there are magnetic-like forces at play that cause signals with the same meaning to be drawn together and signals with different meanings to be drawn apart. In the real-time visualisation of the simulation tool, the bubbles representing vowels can be observed to act in such a way. More specifically, there are two things that appear prominently: bubbles of the same colour are attracted to each other, forming groups of similarly coloured bubbles. Differently coloured bubble groups repel each other. Since attraction and repulsion are dominant in this description, it is easy to talk about the bubbles as having magnetic properties. The reason why this happens has been the topic of this thesis.

In observing the visualisation, sometimes several groups of bubbles (for example three) move in on one group, after which that latter group dissolves completely. In that case, the whole latter group becomes so unsuccessful in aligning games that it needs to go. In the simulation, a new group of bubbles of the lost colour will reappear, since agents will create associations for referents they have no associations for. In reality, events like this can happen as well. What happens after the disappearance of a shared association depends on the importance of the meaning the signal refers to; it could be that a meaning is outdated and that it is not reintroduced with a new signal after dying out. The DEViL is not capable yet of modelling these type of events, since semantics are reduced to almost nothing. When a sufficiently realistic semantic layer is added though, simulating this is possible and really interesting.

Something else worth mentioning is that under certain parameter settings, extended chain shifts can be observed frequently. It is interesting to see that given alignment as the only source of vowel change, chain shifts already belong to the possible behavioural outcomes. Of course, these shifts' directions are of a more random nature than Labov's principles prescribe, but future adjustments to the model such as the addition of co-articulation as a source of directed change are already a big step in the direction of realistic vowel chain shifts.

There are several ways in which simple extensions to this model open doors to a wide, new range of hot and interesting research directions. In my current opinion, three specific extensions are particularly worthwhile. These are (1) to extend the signal model from vowels to at least syllables, (2) to split up the internal repertoires used for perception and production purposes and (3) to let agents have offspring and die. I will describe these in detail, and point to a few other potential additions to the model.

The first extension is related to the signal space. Currently, the agents use vowels to communicate. First of all, this is not realistic, because humans also utilise consonants to create syllables and larger linguistic structures. Adding this possibility to the agent's signal space makes the model much more humanlike. With it come much more possibilities. One important one is that when agents are able to create syllables, one can test another hypothesis regarding sources of linguistic change: the co-articulation model. If the influence of consonantal context on the realisation of vowels is taken into account, it becomes possible to test whether misperception of co-articulation can drive vowel change. The resulting linguistic dynamics can then be compared with vowel shifts seen in the real world, the prediction being that they are more alike (more according to the principles of change) than the dynamics resulting purely from alignment as a source of change. Co-articulation gives direction to the change, it being biased according to the nature of misperception, in contrast to the random direction of change induced by alignment.

The second extension further benefits the possibility of researching co-articulation. Currently, agents use the same lexicon for both perceiving and producing vowels. In reality, these two processes involve completely different mechanics, making this approach unrealistic. For perception, frequencies are analysed to recognise prototype sounds, whereas for production, the tongue and other speech organs are controlled to create prototype utterances. If the model reflects this division of labour, phenomena coming forth from either perception or production specifically can be tested. These include the classic idea that languages are attracted to a state in which there exists minimal articulatory effort and maximal perceptual distinctiveness, but also the more recent idea of co-articulation as a source of change.

The third extension also adds to the model's validity. Currently, an agent is born on the first time step of the simulation and continues to live for the rest of it. A real human, like every object, is subject to physical breakdown and will eventually pass away. In addition, humans produce offspring through intercourse. Modelling human life is beneficial in various ways. For instance, it introduces new agents (children) which need to acquire the language. Since children are still flexible, this acquisition is a potential source of linguistic change. Labov (1994) shows this in several studies of change in real and apparent time. The possibility of analysing the population for change in apparent time is actually another benefit. It is also a step towards the inclusion to the model of biological evolution.

Of course there are many more cool potential extensions. Here are a few things that would be interesting to study in the near future:

- implement types of illocutionary speech acts (Searle, 1971),
- implement different aligning strategies such as divergence (Giles et al., 1991),
- model change from above as opposed to change from below,
- recreate dialect continua by expanding the number of locations,
- recreate the Great Vowel Shift,
- simulate the evolution of Indo-European languages from Proto-Indo-European to now, and
- incorporate meaning shifts in the model.

Chapter 11

Conclusion

There were three goals for this thesis, which have all been successfully reached. A model of linguistic dynamics was developed, which was used to posit three hypotheses regarding the phenomena of linguistic formation, change, and divergence.

The model is both compact and capable of explaining several phenomena that are important in the fields of (evolutionary) linguistics. It explains lexical formation as being driven by learning and forgetting. Thereby, it renders the view that synonym punishment is crucial for lexical formation has to move more to the background. The model explains lexical change and divergence as being driven by alignment (among possible other sources) and geographical constraints respectively. The fact that the link between alignment and change is still under established is at least remarkable.

Furthermore, a simulation tool (the DEViL) was designed and implemented, in which agents interact by playing so-called aligning games with vowels as signals. The aligning game is a new type of language game based on the guessing game that lets agents update their lexicons in a manner that is truthful to the proposed model of linguistic dynamics. These agents are situated on a map of locations, through which they can travel. With the tool, one can simulate all kinds of linguistic phenomena relating to lexical formation, vowel change, and vowel divergence. In this study, it was used to provide support all three hypotheses based on the model of linguistic dynamics.

The combination of the model and the tool give a good perspective for future studies. By applying different expansions, various hot and interesting topics can be researched, ranging from the simulation of Labov's (1994) vowel change from above to testing Ohala's (1993) hypothesis that misperception of co-articulation is a source of change.

Bibliography

Bickerton, D. (1975). *Dynamics of a Creole System.* Cambridge University Press, Cambridge.

Borensztajn, G. (2006). Luc steels, the talking heads experiment and cognitive philosophy. A tutorial accompanying the presentation in Current Issues, Amsterdam.

Bybee, J. L. (2006). *Frequency of Use and the Organization of Language.* Oxford University Press, Oxford.

Cangelosi, A. and Parisi, D. (2002). Computer simulation: a new scientific approach to the study of language evolution. In Cangelosi, A. and Parisi, D., editors, *Simulating the Evolution of Language*, chapter 1, pages 3–28. Springer Verlag, London.

Christiansen, M. H. and Kirby, S. (2003). Language evolution: the hardest problem in science? In Christiansen, M. H. and Kirby, S., editors, *Language Evolution*, chapter 1, pages 1–15. Oxford University Press, Oxford.

Croft, W. (2000). *Explaining language change : an evolutionary approach.* Pearson Education, Essex.

Crothers, J. (1978). Typology and universals of vowel systems in phonology. In Greenberg, J. H., editor, *Universals of human language*, volume 2, pages 95–152. Stanford University Press, Stanford, CA.

de Boer, B. (2001). *The Origins of Vowel Systems: studies in the evolution of language.* Oxford University Press, Oxford.

de Boer, B. (2010). The perceptual effect of air sacs. In Smith, A. D. M., Schouwstra, M., de Boer, B., and Smith, K., editors, *The Evolution of Language: Proceedings of the 8th International Conference*, pages 395–396. World Scientific.

de Boer, B. and Zuidema, W. (2010). Models of language evolution: Does the math add up? In de Boer, B. and Zuidema, W., editors, *Models of Language Evolution: Does the Math Add Up?*, Utrecht.

Dowell, S. (2006). Investigating the vowel systems of contact languages using a population of artificial agents.

Giles, H., Coupland, J., and Coupland, N. (1991). Accomodation theory: Communication, context and consequences. In Giles, H., Coupland, J., and Coupland, N., editors, *Contexts of accommodation*, pages 1–68. Cambride University Press, Cambridge.

Giles, H. and Ogay, T. (2006). Communication accomodation theory. In Whaley, B. B. and Samter, W., editors, *Explaining communication: contemporary theories and exemplars*, pages 325–344. Lawrence Erlbaum, Mahwah, NJ.

Gong, T. (2010). A brief overview of computational modeling in evolutionary linguistics. In de Boer, B. and Zuidema, W., editors, *Models of Language Evolution: Does the Math Add Up?*, Utrecht.

Gray, R. D. and Atkinson, Q. D. (2003). Language-tree divergence times support the anatolian theory of indo-european origin. *Nature*, 426:435–439.

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

Harrington, J. (2000). Does the queen speak the queen's english? *Nature*, 408:927–928.

Harrington, J. (2005). An acoustic analysis of happy-tensing in the queen's christmas broadcasts. *Journal of Phonetics*, 34:439–457.

Hira, Z. M. (2009). Simulating creole and language formation.

Hira, Z. M. and Bartlett, M. (2010). Simulating creole and dialect formation. In Smith, A. D. M., Schouwstra, M., de Boer, B., and Smith, K., editors, *The Evolution of Language: Proceedings of the 8th International Conference*, pages 419–420. World Scientific.

Hombert, J.-M. (2010). Ecological and sexual explanations for larynx lowering. In Smith, A. D. M., Schouwstra, M., de Boer, B., and Smith, K., editors, *The Evolution of Language: Proceedings of the 8th International Conference*, pages 421–422. World Scientific.

Hurford, J. (2002). Expression/induction models of language evolution: dimensions and issues. In Briscoe, T., editor, *Linguistic Evolution through Language Acquisition: Formal and Computational Models*, chapter 10. Cambridge University Press.

Ke, J.-Y. and Holland, J. H. (2006). Language origin from an emergentist perspective. *Applied Linguistics*, 27(4):691–716.

Kirby, S. (2002). Natural language from artificial life. *Artificial Life*, 8:185–215.

Kirby, S., Cornish, H., and Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, 105(31):10681–10686.

Kleber, F., Harrington, J., and Reubold, U. (2010). The relationship between the perception and production of coarticulation during a sound change in progress. to appear.

Kwisthout, J., Vogt, P., Haselager, P., and Dijkstra, T. (2008). Joint attention and language evolution. *Connection Science*, 20(4–3):155–171.

Labov, W. (1994). *Principles of Linguistic Change. Volume 1: Internal Factors*. Blackwell, Oxford.

Lewis, D. K. (1969). *Convention: A Philosophical Study*. Harvard University Press, Cambridge, MA.

Lieberman, P. (2006). *Toward an Evolutionary Biology of Language*. Belknap Press of Harvard University Press, Cambridge, MA.

Liszka, J. J. (1996). *A General Introduction to the Semeiotic of C.S. Peirce*. Indiana University Press, Bloomington.

Livingstone, D. (2002). The evolution of dialect diversity. In Cangelosi, A. and Parisi, D., editors, *Simulating the Evolution of Language*. Springer Verlag, London.

Livingstone, D. and Fyfe, C. (2000). Modelling language-physiology coevolution. In Knight, C., Hurford, J. R., and Studdert-Kennedy, M., editors, *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*, Cambridge University Press.

Maddieson, I. (1984). *Patterns of Sounds*. Cambridge University Press, Cambridge.

Martinet, A. (1955). *Economie des changements phonétiques*. Francke, Berne.

Ohala, J. J. (1993). Sound change as nature's speech perception experiment. *Speech Communication*, 13:155–161.

Pickering, M. J. (2006). The dance of dialogue. *The Psychologist*, 19(12).

Pickering, M. J. and Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(2):169–190.

Schepens, J., Dijkstra, T., and Grootjen, F. (2010). Distributions of cognates in europe as based on levenshtein distance. to appear.

Searle, J. R. (1971). *The Philosophy of Language*. Oxford University Press, Oxford.

Senghas, A., Ozyurek, A., and Goldin-Meadow, S. (2010). The evolution of segmentation and sequencing: Evidence from homesign and nicaraguan sign language. In Smith, A. D. M., Schouwstra, M., de Boer, B., and Smith, K., editors, *The Evolution of Language: Proceedings of the 8th International Conference*, pages 279–288. World Scientific.

Steels, L. (1999). *The Talking Heads experiment*. Laboratorium, Antwerpen. Limited Pre-edition.

The Five Graces Group (2010). Language is a complex adaptive system. In Ellis, N. C. and Larsen-Freeman, D., editors, *Language as a Complex Adaptive System*, The Best of Language Learning Series, chapter 1. Wiley-Blackwell, Malden.

Vogt, P. (2003). Thsim v3.2: The talking heads simulation tool. In Banzhaf, W., Christaller, T., Dittrich, P., Kim, J. T., and Ziegler, J., editors, *Proceedings of the 7th European Conference on Artificial Life, ECAL 2003*. Springer-Verlag.

Appendices

Appendix A

Java code

In this appendix, the most relevant Java code is presented. First, the aligning game is discussed, after which we go through some of the measures used in the simulations. An aligning game starts with an agent initiating a conversation. If a random number is below the interaction rate, the agent picks a partner and starts a conversation.

Agent code: initiating a conversation

```
public void initiateConversation() {
        HighQualityRandom generator = new HighQualityRandom();
        double r = generator.nextDouble();

        if (r < Parameters.interactionRate) {
                Agent partner = this.wanderForConversation();
                if (partner != null) {
                        Conversation myConv = new Conversation(this, partner);
                }
        }
}
```

The conversation protocol, in this case that of the aligning game, then takes over. It asks the two agents, the speaker and the hearer, to act according to the rules of the aligning game. That is, the speaker chooses a topic, gives a hint and the guesser tries to guess what the topic is. Depending on the guess, the agents update their lexicons.

Conversation code: controlling the alignment game

```
public Conversation(Agent a1, Agent a2) {
        speaker = a1;
        hearer = a2;

        Referent topic = speaker.chooseTopic();
        Signal hint = speaker.giveHint(topic);
        Referent guess = hearer.makeGuess(hint);

        if (guess == null) {
                hearer.takeOver(topic, hint);
        } else {
                if (guess != topic) {
```

```
                hearer.takeOver(topic, hint);
            } else {
                speaker.reward();
                hearer.rewardAndAdjust(hint);
            }
        }
}
```

Agents generally use information from their location, their lexicon and memory to fulfill the tasks of choosing a topic, giving a hint for a topic and making a guess given a hint. For example, the Lexicon class is able to retrieve the best available guess, given a signal (a vowel). The agent then only has to return this guess (or point at it). In the GiveHint function, it can be seen that Gaussian noise is added to the utterance.

Agent code: alignment game actions

```
public Referent chooseTopic() {
        return this.location.getRandomTopic();
}

public Signal giveHint(Referent topic) {
        LexicalElement template = lexicon.getElementFor(topic);
        if (template == null) {
                this.lexicon.addElement(topic);
                template = lexicon.getElementFor(topic);
        }
        memory = template;
        Signal hint = template.getSignal().withAddedGaussianNoise();
        return hint;
}

public Referent makeGuess(Signal hint) {
        LexicalElement template = lexicon.getBestGuess(hint);
        if (template != null) {
                memory = template;
                return template.getReferent();
        }
        return null;
}
```

After the aligning game, the agents update their lexicons. As can be seen below, agents update their lexicon positively ("pos"). This is implemented like this because it is also possible to run simulations in which agents punish associations, for example after unsuccessful games or using synonym punishment. Because these update processes were not used, the code for it is not printed in this appendix. Note that when taking over an association, an agent actually checks if it already knows the association. This seems superfluous, but under certain circumstances, the hearer can misguess although it knew the hint-topic relation, for example when it has synonyms.

Agent code: updating processes

```
public void reward() {
        this.lexicon.update(memory.getReferent(), memory.getSignal(), "pos");
}
```

```
public void rewardAndAdjust(Signal hint) {
        this.lexicon.update(memory.getReferent(), memory.getSignal(), "pos");

        double[] adjustment = new double[Signal.getNumberOfFeatures()];
        for (int fea = 0; fea < Signal.getNumberOfFeatures(); fea++) {
                double memfea = memory.getSignal().getFeature(fea);
                double hintfea = hint.getFeature(fea);
                if (memfea - hintfea > Parameters.adjustmentRate) {
                        adjustment[fea] = -Parameters.adjustmentRate;
                } else if (hintfea - memfea > Parameters.adjustmentRate) {
                        adjustment[fea] = Parameters.adjustmentRate;
                } else {
                        adjustment[fea] = 0;
                }
        }
        lexicon.adjust(memory.getReferent(), memory.getSignal(), adjustment);
}

public void takeOver(Referent topic, Signal hint) {
        if (lexicon.elementExists(topic, hint)) {
                lexicon.update(topic, hint, "pos");
        } else {
                lexicon.addElement(topic, hint);
        }
}
```

Regarding the computation of the lexical distance between locations, in Chapter 5 it was explained that it consisted of three measures: a vowel space based distance, a synonymy based distance and an associative strength based distance. These measures were combined in a 3:1:1 ratio. In the Java code, the first measure was named LAF, the second EL and the third WE. The following code was used to compute LAF, which was the most involved computation. Note that 112 is the biggest distance there can be between two vowels in the vowel space.

Location code: computing vowel space based distance

```
public double computeLAF() {
        double measure_laf = 0;
        for(Agent agA : inhabitants) {
        for(Agent agB : inhabitants) {
        if(agA!=agB) {

                for(int refID = 0; refID < numberOfReferents; refID++) {
                        ArrayList<LexicalElement> siA =
                                agA.getLexicon().getAllSignalsFor(referents.get(refID));
                        ArrayList<LexicalElement> siB =
                                agB.getLexicon().getAllSignalsFor(referents.get(refID));
                        double expect_we = 0;
                        double expect_laf = 0;
                        for(LexicalElement el : siA) {
                                double w = el.getWeight();
                                expect_we += w;
                                double smallest = 112;
                                for(LexicalElement elOther : siB) {
                                        double contender = el.getSignal().
                                                distanceTo(elOther.getSignal());
                                        if(contender < smallest)
                                                smallest = contender;
                                }
                                expect_laf += w * smallest / 112;
```

```
                        }
                        for(LexicalElement el : siB) {
                                double w = el.getWeight();
                                expect_we += w;
                                double smallest = 112;
                                for(LexicalElement elOther : siA) {
                                        double contender = el.getSignal().
                                                distanceTo(elOther.getSignal());
                                        if(contender < smallest)
                                                smallest = contender;
                                }
                                expect_laf += w * smallest / 112;
                        }

                        if(expect_we > 0)
                                measure_laf += expect_laf / expect_we;
                }
        }
        }
        }
        return 100* measure_laf / (numberOfReferents * inhabitants.size()
                * inhabitants.size()-1);
}
```

Below is the code for the other two measures. Some of the work for these computations has also been done in the previous computation, which makes the code a little redundant. In future versions, this will be better maintained (and documented).

```
Location code: computing vowel space based distance
public double computeEL() {
        double measure_el = 0;
        for(Agent agA : inhabitants) {
        for(Agent agB : inhabitants) {
        if(agA!=agB) {

                for(int refID = 0; refID < numberOfReferents; refID++) {
                        ArrayList<LexicalElement> siA =
                                agA.getLexicon().getAllSignalsFor(referents.get(refID));
                        ArrayList<LexicalElement> siB =
                                agB.getLexicon().getAllSignalsFor(referents.get(refID));
                        double expect_el = ((double)siA.size() + (double)siB.size()) / 2;

                        if(expect_el > 0)
                                measure_el += Math.abs((double)siA.size() - expect_el) /
                                        expect_el;
                }
        }
        }
        }
        return 100* measure_el / (numberOfReferents * inhabitants.size() *
                inhabitants.size()-1);
}

public double computeWE() {
        double measure_we = 0;
        for(Agent agA : inhabitants) {
        for(Agent agB : inhabitants) {
        if(agA!=agB) {

                for(int refID = 0; refID < numberOfReferents; refID++) {
                        ArrayList<LexicalElement> siA =
```

```
                        agA.getLexicon().getAllSignalsFor(referents.get(refID));
                ArrayList<LexicalElement> siB =
                        agB.getLexicon().getAllSignalsFor(referents.get(refID));
                double expect_we = 0;
                for(LexicalElement el : siA) {
                        expect_we += el.getWeight();
                }
                double we0 = expect_we;
                for(LexicalElement el : siB) {
                        expect_we += el.getWeight();
                }
                expect_we /=2;

                if(expect_we > 0)
                        measure_we += Math.abs(we0 - expect_we) / expect_we;
            }
        }
        }
        }
        return 100* measure_we / (numberOfReferents * inhabitants.size() *
                inhabitants.size()-1);
}
```

Appendix B

List of abbreviations

$\boldsymbol{A}$	Number/set of agents, a denotes one agent
$\boldsymbol{\aleph}$	Number/set of associations, α denotes one association
$\boldsymbol{c}$	Communicative success in percentage of successful games
CAS	Complex Adaptive System
CAT	Communication Accomodation Theory
$\boldsymbol{\delta}$	Rate of change in reward δ_r, decay δ_d or alignment δ_a
E/I	Expression/Induction
$\boldsymbol{HLD}$	Historical Lexical Distance
IAM	Interactive Alignment Model
$\boldsymbol{ILD}$	Interlocal Lexical Distance
$\boldsymbol{ILD_{int}}$	Boundary in ILD of mutual intelligibility
ILM	Iterated Learning Model
IPA	International Phonetic Alphabet
$\boldsymbol{L}$	Number/set of locations, l denotes one location
$\boldsymbol{LD_{rand}}$	Expected ILD of two randomly picked, independent languages
$\boldsymbol{LV}$	Local variance
$\boldsymbol{\omega}$	Strength of geographical constraint ω_g or association ω_a, which is initiated at ω_0 and is bounded by ω_{min} and ω_{max}
$\boldsymbol{\Psi_{ac}}$	Acoustic noise

$\boldsymbol{R}$	Number/set of referents, r denotes one referent
RP	Received Pronunciation
$\boldsymbol{s}$	Synonymy
SSB	Standard Southern-British
$\boldsymbol{\tau}$	Travel rate
$\textbf{UPSID}_{451}$	UCLA Phonological Segment Inventory Database