# CALGARY WORKING PAPERS IN LINGUISTICS

Number 24 Fall 2002

**THE**
**UNIVERSITY**
**OF**
**CALGARY**

Department of Linguistics
The University of Calgary
2500 University Drive N.W.
Calgary, Alberta, CANADA
T2N 1N4

## FOREWORD

The editors of this volume, Jodi Edwards and Michael Dobrovolsky are pleased to present the twenty-fourth issue of the *Calgary Working Papers in Linguistics* published by the Department of Linguistics at the University of Calgary. The papers contained in this volume represent works in progress and as such should not be considered in any way final or definitive.

This issue of *CWPL* includes papers from both graduate students and professors in the Department of Linguistics. This year, we are also pleased to include a special section devoted to the undergraduate honours theses of our second year MA candidates, who all completed undergraduate degrees in this department. The articles in this journal discuss a range of topics from the fields of syntax, phonetics and phonology.

The first paper, by Martha McGinnis, investigates the aspectual properties of phrasal idioms, offering support for the theory of Distributed Morphology. Ilana Mezhevich's paper discusses the argument status of postverbal DPs in English resultative constructions, English verb-particle constructions and Russian prefixed verbs. In our special undergraduate theses section, Rebecca Hanson presents a case study for the acquisition of English onsets and Jennifer Mah investigates the role of explicit linguistic knowledge in L2 phonological acquisition. Jodi Edwards' paper investigates the neural substrates of phonological processing and Sheena Van Der Mark presents data analyzing the acoustic correlates of Blackfoot prominence.

We wish to express our sincere gratitude to Linda Toth for her assistance in this project. We would also like to thank the University of Calgary Department of Linguistics and Graduate Students' Association for providing the necessary funding to produce this volume. A final word of thanks is owed to each of our contributors for their submissions to *CWPL*, Volume 24.

 ISSN 0823-0579.

## CALL FOR PAPERS

*Calgary Working Papers in Linguistics* is an annual journal which includes papers by faculty and students in Linguistics and related disciplines, both at the University of Calgary and elsewhere.

The editors would like to encourage all readers to submit papers for future publication. The deadline for submission of papers is **August 31, 2003** in order to meet an autumn publication date. The editors would like contributions on 3 1/2" Micro Floppy Disks (preferably formatted for Microsoft Word for Macintosh version 5 or higher). We further request that the submissions follow the Style Sheet provided at the end of the journal. All submissions should be camera-ready. Page numbers should not be included on the front of the papers, but should be lightly printed on the back of the pages in pencil. Authors should submit their papers to the address listed below. The editors reserve the right to return papers for revisions if they do not conform to the Style Sheet as outlined at the end of the journal. Appearance of papers in this volume does not preclude their publication in another form elsewhere.

Any correspondence should be sent to the address below:

Editors, CWPL
Department of Linguistics
The University of Calgary
2500 University Drive N.W.
Calgary, Alberta, Canada
T2N 1N4

Phone: (403) 220-5469
Fax: (403) 282-3880

The editor can also be reached by e-mail at the following address: cwpl@ucalgary.ca or toth@ucalgary.ca. Any queries regarding the formatting of papers can also be directed to that address.

The journal is available on a reciprocal exchange basis. If you publish a journal or newsletter which you would like to send us, we will send you our journal exchange gratis. Yearly subscriptions are also available for the following rates: in Canada $10, in the US for $11 and overseas for $12. All prices (including postage) should be remitted in Canadian funds.

To request back issues or to receive information on the contents of back issues, please send a self-addressed stamped (Canadian) envelope to the above address. If requesting this information outside of Canada, please include $2.00 Canadian to cover postage. For those who have access to e-mail, any inquiries may be made to the above e-mail address, thus eliminating the postal costs and ensuring a more timely response to your request.

## Table of Contents

**On A Systematic Component of Meaning in Idioms**

Martha McGinnis
University of Calgary

**Abstract**

It has traditionally been assumed that the meaning of some or all phrasal idioms is non-compositional. However, I argue here that the aspectual meaning of idioms is completely systematic: there are no special aspectual restrictions on idioms, and moreover, the aspectual properties of an idiom are compositional, combining the aspectual properties of its syntactic constituents in the usual way. I show that this observation supports the theory of Distributed Morphology (Halle & Marantz 1994).

## 1 Aspectual Classes of Idioms

It is worth noting that all aspectual classes contain idiomatic VPs. In what follows, I will assume the familiar Vendlerian classes (states, activities, achievements, and accomplishments), identified by an array of tests from the literature (see Vendler 1967, Dowty 1979, Mittwoch 1991, among many others). Subclasses of achievements and accomplishments will also be distinguished. However, the aspectual parallelism between idiomatic and non-idiomatic VPs is independent of this classification.

States and activities are atelic predicates, which can be modified by adverbial PPs with *for*, but not by adverbial PPs with *in*, at least not with the sense that the state of affairs denoted by the VP ends in the time specified:

| | | | |
|---|---|---|---|
| (1) | a. | Harry knew the truth for years/#in an hour. | ATELIC |
| | b. | Hermione pushed the cart for an hour/#in an hour. | ATELIC |

The crosshatch (#) indicates the availability of an alternative reading. For instance, the examples with *in*-phrases in (1) are marginally acceptable on the interpretation that the state of affairs denoted by the VP begins, rather than ends, when an hour has elapsed.

In English, states and activities can be distinguished using the progressive: states generally cannot occur in the progressive, while activities can:

| | | | |
|---|---|---|---|
| (2) | a. | *Harry is knowing the truth. | STATE |
| | b. | Hermione is pushing the cart. | ACTIVITY |

The same classes can be identified in idiomatic VPs. The idiomatic state *be the cat's pyjamas* ("be terrific") can occur with a *for*-phrase, but not with an *in*-phrase (3a) — except on the marginal reading noted above — or with the progressive (3b).

(3) a. Hermione was the cat's pyjamas for years/#in an hour. STATE
b. * Hermione is being the cat's pyjamas.

On the other hand, the idiomatic activity *jump through hoops* ("try to meet exacting expectations") can occur with a *for*-phrase and the progressive, but not with an *in*-phrase:

(4) a. Harry jumped through hoops for years/#in an hour. ACTIVITY
b. Harry is jumping through hoops.

Unlike states and activities, accomplishments and achievements are telic: they allow modification by *in*-phrases. (5a) is true if Harry finished climbing the mountain within an hour after he started. The event in (5b) both begins and ends in an instant.

(5) a. Harry climbed the mountain in an hour. ACCOMPLISHMENT
b. Hermione noticed the painting in an instant. ACHIEVEMENT

Several tests have been used to distinguish achievements from accomplishments. For example, accomplishments (6a), but not achievements (6b), generally allow modification by a *for*-phrase. Moreover, accomplishments can occur in the progressive (7a), while achievements generally cannot (7b).

(6) a. Harry climbed the mountain for an hour. ACCOMPLISHMENT
b. # Hermione noticed the painting for an hour. ACHIEVEMENT

(7) a. Harry was climbing the mountain. ACCOMPLISHMENT
b. # Hermione was noticing the painting. ACHIEVEMENT

The examples in (6b) and (7b) may marginally allow an iterative reading, in which Hermione kept noticing the painting again and again.

Another difference is that accomplishments, but not achievements, can be halted in midstream. If VP is an achievement, then *X stopped VPing* entails that *X VPed*. If VP is an accomplishment, this entailment does not hold: instead, *X stopped VPing* can mean that the event stopped before it was completed. For example, (8a) could mean that Harry did not climb the mountain, while (8b) entails that Hermione noticed the painting. Moreover, if VP is an achievement, *X stopped VPing* carries an iterative implicature — for example, (8b) suggests that Hermione noticed the painting several times.

(8) a. Harry stopped climbing the mountain. ACCOMPLISHMENT
b. # Hermione stopped noticing the painting. ACHIEVEMENT

Idiomatic VPs also show the characteristics of accomplishments and achievements. For example, the idiomatic accomplishment *pay one's dues* ("earn one's right to something") can be modified by an *in*-phrase (9a), as can the idiomatic achievement *strike paydirt* ("gain something valuable") (9b).

(9) a. Hermione paid her dues in ten years. ACCOMPLISHMENT
b. Harry struck paydirt in an hour. ACHIEVEMENT

Idiomatic achievements and idiomatic accomplishments can also be distinguished from each other, as illustrated in (10)-(11). I leave the details for the reader to verify.

(10) a. Hermione paid her dues for ten years. ACCOMPLISHMENT
b. Hermione was paying her dues.
c. Hermione stopped paying her dues.
(11) a. #Harry struck paydirt for an hour. ACHIEVEMENT
b. # Harry was striking paydirt.
c. #Harry stopped striking paydirt.

Actually, the three tests just given to distinguish achievements from accomplishments do not yield exactly the same results. For the sake of exposition, let us assume that the *X stopped VPing* test is diagnostic of the split between accomplishments and achievements. So stated, there is a subclass of accomplishments that cannot be modified by a *for*-phrase, and a subclass of achievements that combines more easily with the progressive. An example of the "*for*-less" accomplishment (FLA) subclass is given in (12). As with other accomplishments, the *X stopped VPing* context does not entail that this accomplishment was completed (12a). Furthermore, this VP allows modification by an *in*-phrase (12b), and can be used in the progressive without an iterative meaning (12c).

(12) a. Hermione stopped burying the treasure.
b. Hermione buried the treasure in five minutes.
c. Hermione was burying the treasure.

However, unlike other accomplishments, this VP does not allow a *for*-phrase:

(13) #Hermione buried the treasure for five minutes. FLA

More precisely, the *for*-phrase cannot modify the burying process, but in some cases can modify the result of this process. For example, (13) would be true if Hermione dug up the treasure five minutes after burying it.

As the reader may confirm, there are idiomatic VPs with the same characteristics as these *for*-less accomplishments, for example *get one's act together* ("get organized") (14). Again, (14d) is fine if Harry got disorganized a month after getting organized.

(14) a. Harry stopped getting his act together. FLA
b. Harry got his act together in one semester.
c. Harry was getting his act together.
d. # Harry got his act together for a month.

The achievement class is also divided into subclasses, one of which is more compatible with the progressive. With this "*prog*-ful" subclass of achievements (PFA), as with other achievements, the *X stopped VPing* context entails that the VP event was completed, with an implicature that it was completed iteratively (15a). This subclass likewise allows modification by an *in*-phrase (15b), but not by a *for*-phrase (15c). Nonetheless, in this subclass the progressive does not imply an iterative reading. Instead, (15d) seems to mean something like "Hermione was searching for the exit."

(15) a. #Hermione stopped finding the exit. PFA
b. Hermione found the exit in ten minutes.
c. #Hermione found the exit for ten minutes.
d. Hermione was finding the exit.

Idiomatic VPs also occur in this subclass, such as the VP *get to first base* ("kiss someone"). Again, the reader may verify the parallels between (15) and (16).

(16) a. #Harry stopped getting to first base. PFA
b. Harry got to first base after one date.
c. #Harry got to first base for one evening.
d. Harry was getting to first base.

To summarize, any aspectual classification of non-idiomatic VPs also applies to idiomatic VPs. In this sense, idiomatic VPs are aspectually systematic.[1] More intriguing is the observation that the aspectual properties of idiomatic VPs are, at least in part, syntactically derived. I turn to this issue now.

---

[1] Additional examples of each class are readily available: states (*have bigger fish to fry, take the cake*), activities (*beat around the bush, push one's luck*), accomplishments (*run X into the ground, climb the ladder of success*), achievements (*drop the ball, kick the bucket*), *for*-less accomplishments (*make a name for oneself, go around the bend*), and *prog*-ful achievements (*hit one's stride, find one's tongue*).

## 2 Aspectual Compositionality in Idioms

The claim that idioms are aspectually compositional bears on a recent debate concerning the correspondences between syntax and meaning. It is generally acknowledged that words are associated with two types of semantic information, which Rappaport Hovav and Levin (1998) call the *structural* and *idiosyncratic* components of meaning. The structural component of meaning interacts with the syntax, while the idiosyncratic component makes fine-grained distinctions that are irrelevant to the syntax. In Jackendoff's theory of Representational Modularity, both types of meaning are encoded at Conceptual Structure (CS); structural meaning is "visible" to correspondence rules between syntax and CS, while idiosyncratic meaning is not (1997:220). By contrast, the theory of Distributed Morphology (Halle & Marantz 1994) maintains that structural components of meaning are bundled into lexical items manipulated by the syntax, while idiosyncratic components are added post-syntactically, by reference to a list known as the Encyclopedia.

These two approaches make different predictions for the interpretation of idioms. Jackendoff argues that idioms are syntactically complex, but differ from non-idioms in the mapping to interpretation. In Representational Modularity terms, the head V of a non-idiomatic VP maps to a lexical conceptual structure (LCS), while its arguments map onto slots in this structure. For example, the LCS of a transitive verb like *kick* would have two argument slots. In the case of an idiomatic VP, however, the whole VP maps to an LCS, while the syntactic arguments of the verb need not map onto argument slots. For example, *kick the bucket* has no slot for *the bucket*: the idiomatic LCS of this VP is the same as the LCS for the intransitive verb *die* (Jackendoff 1997:169). In short, Representational Modularity treats idioms as involving an arbitrary mapping between CS and syntactic structure. Since this theory encodes both structural and idiosyncratic meaning at CS, both types of meaning are predicted to be subject to arbitrary mapping.

In Distributed Morphology, however, the structural components of meaning are assembled in the syntax. This theory predicts that the syntactic derivation of idioms has semantic consequences. Marantz (1997) suggests that one such consequence is aspectual. He argues that *kick the bucket* cannot mean "die", because it has the punctual aspect of a transitive VP with a definite complement. Thus, though (17a) is fine, (17b) is out.

(17) a. Hermione was dying for weeks.
b. * Hermione was kicking the bucket for weeks.

If this analysis is correct, it predicts that even if a VP has a non-compositional idiosyncratic meaning, it will have a compositional structural meaning. Specifically, it will have the same aspectual properties as any VP with the same syntactic properties.

One reason to suppose that aspect is a structural component of meaning is that it interacts with structural properties of the sentence (see Tenny 1987, among others). For example, when the verb *eat* takes a DP complement, the VP is generally telic, allowing *in*-phrase modification and disallowing *for*-phrase modification (18a). When it takes no complement, the VP is atelic, disallowing *in*-phrase modification, and allowing *for*-phrase modification (18b).

(18) a. Hermione ate her vitamins {in two seconds flat/*for five minutes}.
b. Harry ate for/*in a week.

The semantic properties that distinguish bare plural and mass DPs from other DPs also seem to be structural components of meaning, since they affect the formal expression of DPs, the choice of determiners, and so forth. When the complement of *eat* is a bare plural or mass DP, the VP has the same atelic aspectual properties as with intransitive *eat*:

(19) Harry ate turkey (sandwiches) for/*in a week.

If idioms have compositional aspect, the structure of an idiom should also have aspectual consequences. This prediction is confirmed. *Eat one's words* ("admit to being wrong") has the telic aspectual properties of the non-idiomatic *eat one's vitamins* (20a), while *eat crow* ("lose one's pride") has the atelic aspectual properties of *eat turkey* (20b).

(20) a. Hermione ate her words {in two seconds flat/*for five minutes}.
b. Harry ate crow for/*in a week.

These facts suggest that, even in idiomatic VPs, the structural component of meaning is not arbitrarily related to the syntax , as Representational Modularity predicts, but instead is derived from it.

This observation also has implications for an account of the passivizability of idioms. It has long been noted that some idioms may passivize, while others cannot (Katz & Postal 1964, Fraser 1970, Katz 1973, Fiengo 1974, Newmeyer 1974). For example, (21a) retains the idiomatic meaning of the active, while (21b) has only a literal meaning.

(21) a. The beans were spilled (by Hermione).
b. # The bucket was kicked (by Hermione).

Nunberg et al. (1994) propose that this difference arises from a distinction between compositional and non-compositional idioms. They argue that cases like (21a) are composed of subparts with idiosyncratic meanings. For example, in *spill the beans*, *spill* takes on a special meaning like "divulge", and *beans* takes on a meaning like "secret". On the other hand, they claim that cases like (21b) are lexically stored as a whole, and thus cannot undergo syntactic operations.

However, the aspectual facts suggest that the structural component of meaning is always compositionally derived from the syntax. Thus, even idiomatic VPs that cannot undergo passivization have compositional aspect. For instance, the VPs in (20) cannot be passivized and retain their idiomatic interpretations:

(22) a. #Her words were eaten (by her/Hermione).
b. # Crow was eaten (by Harry).

Moreover, *kick the bucket* (an achievement) and *saw logs* (an activity), which cannot passivize, are aspectually identical to their non-idiomatic counterparts, except that an iterative reading of the idiomatic *kick the bucket* is pragmatically unavailable. The non-idiomatic *kick the hand-grenade*, which also disfavors this reading, is completely parallel to the idiom. Thus the availability of passivization cannot be tied to a distinction between compositional and non-compositional idioms. One alternative worth exploring is that an idiom is passivizable if its idiosyncratic meaning is assigned to a thematic representation, but not if it is assigned to a morphosyntactic representation (Lebeaux 1988).

The facts presented above demonstrate that the meaning of idioms is not entirely arbitrary: the structural component of meaning (specifically, aspect) is both systematic and compositional. This observation supports the claim of Distributed Morphology that structural meaning, but not idiosyncratic meaning, is built in the syntax.

**Acknowledgements**
Thanks go to Heidi Harley and Alec Marantz for useful discussion of this work.

**References**

Dowty, David R. 1979. *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ.* Dordrecht: Reidel.

Fiengo, Robert. 1974. Semantic conditions on surface structure. Doctoral dissertation, MIT, Cambridge, Mass.

Fraser, Bruce. 1970. Idioms within a transformational grammar. *Foundations of Language* 6:22–42.

Halle, Morris & Alec Marantz. 1994. Distributed Morphology and the pieces of inflection. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. Kenneth Hale & Samuel Jay Keyser, 111–76. Cambridge, Mass.: MIT Press.

Jackendoff, Ray. 1997. *The architecture of the language faculty.* Cambridge, Mass.: MIT Press.

Katz, Jerrold J. 1973. Compositionality, idiomaticity, and lexical substitution. In *A Festschrift for Morris Halle*, ed. Stephen R. Anderson & Paul Kiparsky, 357–76. New York: Holt, Rinehart, and Winston.

Katz, Jerrold J., and Paul Postal. 1964. *An integrated theory of linguistic descriptions*. Cambridge, Mass.: MIT Press.

Lebeaux, David. 1988. Language acquisition and the form of the grammar. Doctoral dissertation, University of Massachusetts, Amherst, Mass.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own Lexicon. In *University of Pennsylvania Working Papers in Linguistics 4:2: Proceedings of the 21st Annual Penn Linguistics Colloquium*, 201–25. Department of Linguistics, University of Pennsylvania, Philadelphia.

Mittwoch, Anita. 1991. In defence of Vendler's achievements. *Belgian Journal of Linguistics* 6:71–85.

Newmeyer, Frederick. 1974. The regularity of idiom behavior. *Lingua* 34:327–42.

Nunberg, Geoffrey, Ivan A. Sag, & Thomas Wasow. 1994. Idioms. *Language* 70:491–538.

Levin, Beth & Malka Rappaport Hovav. 1998. Building verb meanings. In *The projection of arguments: Lexical and compositional factors*, ed. Miriam Butt & Wilhelm Geuder, 97–134. Stanford, Calif.: CSLI Publications.

Tenny, Carol Lee. 1987. Grammaticalizing aspect and affectedness. Doctoral dissertation, MIT, Cambridge, Mass.

Vendler, Zeno. 1967. *Linguistics in philosophy*. Ithaca: Cornell University Press.

Martha McGinnis
Department of Linguistics
University of Calgary
2500 University Drive NW
Calgary, AB T2N 1N4
Fax: (403) 282-3880
mcginnis@ucalgary.ca

## Resultatives, Particles, Prefixes and Argument Structure

**Ilana Mezhevich**
**University of Calgary**

### 1.0 Introduction

In this paper, I discuss the argument status of postverbal DPs in English resultative constructions, English verb-particle constructions and Russian prefixed verbs. The examples of the three constructions are given in (1): (1a) is a resultative construction, (1b) is a verb-particle construction and (1c) is a Russian prefixed verb:

(1) a. Jane wiped the table clean.
    b. Jane wiped the table down.
    c. Masha vy-terla stol.
       Masha VY-wiped.PF table-ACC
       'Masha wiped the table (clean).'

A resultative construction consists of the main verb that denotes an activity and a resultative phrase that is predicated of the postverbal DP and denotes the state achieved by this DP as a result of the activity denoted by the main verb. Thus, (1a) means that the table became clean as a result of wiping; it can never mean that Jane became clean as a result of wiping.[1] The addition of a resultative phrase turns an activity into an accomplishment by specifying a resulting state.

As was suggested in various studies, the verb-particle construction is a relative of the resultative construction (Levin and Rappaport Hovav 1995; Spencer and Zaretskaya 1998, among others). The base verb in a verb-particle construction is typically an activity and the addition of a particle just like the addition of a resultative phrase often turns this activity into an accomplishment. Thus, in (1b), *wipe* is an activity verb, whereas *wipe down* is an accomplishment.

Russian has neither resultative constructions nor verb-particle constructions. However, it has a process of lexical prefixation that derives new

[1] Rappaport Hovav and Levin (2001) reject the direct object restriction, i.e. the fact that the resulative phrase in a resultative construction is always predicated of the object, on the basis of examples such as *The wise men followed the star out of Bethlehem*, where the resultative phrase expressed as a PP *out of Bethlehem* is predicated of the subject *the wise men* and not the object *the star*. Without going into much detail, for the purposes of this paper, I assume the direct object restriction to be a valid generalization for resultative constructions whith a resultative phrase expressed as an AP.

verbs by adding a prefix to the original base. As Spencer and Zaretskaya (1998) point out, the process of prefixation in Russian almost always changes aspectual properties of the stem, frequently by adding an end point. In (1c) above, the verb stem *teret'* wipe is an acitivity verb, whereas the addition of the prefix *vy-* turns it into an accomplishment and the expression means that the table was wiped, presumably with the intention to make it clean.

One of the properties shared by the three constructions is that they can take direct internal arguments different from those selected by their base verbs, as illustrated by the following examples:

(2) a. We danced our feet black and blue. (Martha McGinnis, p.c.)
    b. *They danced their feet.

(3) a. They blew the bridge up.
    b. *They blew the bridge.

(4) a. Masha is-pisala svoju ruchku.
    Masha IZ-wrote[2] her-ACC pen-ACC
    'Masha's pen has run out of ink'
    b. *Masha pisala svoju ruchku.
    Masha wrote her-ACC pen-ACC
    'Masha wrote her pen.'

In (2), *dance* is an intransitive verb which does not allow a direct object in isolation, as shown by the ungrammaticality of (2b).[3] Similarly, in (3), *blow* does not take *bridge* as its internal argument outside of a verb-particle construction. Finally, in (4), the base verb *pisat'* 'write' does not take *ruchka* 'pen' as its internal argument which explains the ungrammaticality of (4b).

As argued by Levin and Rappaport Hovav (1995), the postverbal DPs in English resultative constructions based on intransitive verbs and unspecified object verbs (e.g. *drink*) are not arguments but *unselected objects*, i.e. objects that the verb

[2] *iz-* is listed in Townsend (1975) and is usually considered as a basic (underlying) form. [z] is devoiced when followed by voiceless consonants.
[3] Intransitive verbs such as *dance* or *laugh* can take cognate objects:

(i) a. Jane and Magnus danced a beautiful dance.
    b. Mary laughed hysterical laughter.

However, for the purposes of the present paper, I do not consider such objects.

does not subcategorize for. On the other hand, Spencer and Zaretskaya (1998) suggest that English resultative constructions, verb-particle constructions and Russian prefixed verbs belong to essentially the same kind of complex predicate. Based on the fact that all three constructions have a very similar semantic structure and that in Russian, DPs that follow prefixed verbs have the properties of a regular direct object they argue that such DPs in all three constructions are arguments of the complex predicate.

I argue that postverbal DPs in English resultative constructions on the one hand and verb-particle constructions and Russian prefixed verbs on the other hand have different argument status. As shown by various studies, English resultatives are syntactically derived constructions (Carrier and Randall 1992; Neeleman and Weerman 1993; Levin and Rappaport Hovav 1995, among others), whereas Russian prefixed verbs are lexically derived (Townsend 1975; Brecht 1985, Zaliznjak and Shmelev 1997, among others). I assume that lexical derivation as opposed to syntactic derivation is less productive and may change a verb's meaning in an unpredictable way. Given this, the addition of a resultative phrase affects neither the meaning nor the argument structure of the main verb (Levin and Rappaport Hovav 1995; Rappaport Hovav and Levin 1998). Therefore, postverbal DPs in resultative constructions based on intransitive verbs are not arguments but unselected objects. In contrast, the addition of a Russian prefix often changes the meaning of the base verb, which in turn may affect its aspectual classification and argument structure. Although the status of English verb-particle constructions is more debatable than that of Russian prefixed verbs, they have more properties of morphological words than syntactic phrases. The addition of a particle, similar to the addition of a prefix, often changes the meaning of the base verb, which in turn may have an effect on its aspectual classification and argument structure. As a result, Russian prefixed verbs and English verb-particle constructions often have different meaning from their base verbs and may have different arguments.

In section 2, I briefly outline syntactic evidence that the verb within a resultative construction has the same meaning and the same syntactic properties as it does in isolation (Levin and Rappaport 1986; Carrier and Randall 1992). It will be shown that middle constructions, adjectival passives and nominalizations are grammatical when derived from resultative constructions based on transitive verbs but unacceptable when derived from resultative constructions based on intransitive verbs. In section 3, I discuss Russian prefixed verbs. I show that Russian prefixed verbs are lexically derived and often have different meaning and different syntactic properties from their original stem (Townsend 1975; Brecht 1985). I also discuss the difference between English resultative constructions and Russian prefixed verbs in terms of compositionality and productivity. Section 4 discusses similarities between Russian prefixed verbs and English verb-particle constructions. I also provide further arguments against analyzing Russian prefixed verbs and English verb-particle constructions as counterparts of English resultatives.

## 2.0 English resultatives – tests for argumenthood

### 2.1 Middle formation

Carrier and Randall (1992:188), following Keyser and Roeper (1984), assume that the rule of middle formation, just like the rule of verbal passive formation suppresses a verb's external argument and its ability to assign accusative case. However, middle formation and verbal passive formation do not apply to the same class of verbs. Some verbs can undergo verb passive formation but not middle formation:

(5) a. The politician was laughed at.
b. *Politicians laugh at easily. (Carrier and Randall (1992:189 (41)).

To account for the contrast between (5a) and (5b), they propose the Argument Structure Condition on middle formation:

(6) Middle formation applies to a verb only if it has a direct internal argument. (Carrier and Randall 1992:189 (42)).

This condition predicts that middle formation can apply to the resultative constructions based on transitive verbs, but not to the resultative constructions based on intransitive verbs. The following examples show that this is indeed the case:

(7) a. She wiped the table clean.
b. This table wipes clean easily.
(Levin and Rappaport Hovav 1995:43 (26a))

(8) a. They drank the teapot dry.
b. *This teapot drinks dry in no time at all.
(Levin and Rappaport Hovav 1995:43 (26b))

## 2.2 Adjectival passive formation

Various studies agree that adjectival passive formation externalizes a direct internal argument (Williams 1981, Levin and Rappaport 1986, Grimshaw 1990, among others). For example, Levin and Rappaport (1986) provide evidence that involves dative verbs such as *feed*, where the adjectival passive formation is possible with their direct internal arguments but ungrammatical with the indirect internal arguments:

(9) a. They fed the baby peas.
b. the fed baby
c. *the fed peas
(Carrier and Randall 1992:193 (53a), (51), and (52) respectively)

According to Levin and Rappaport (1986), *feed* can also have an argument structure where it has two obligatory arguments – theme and goal – and where the direct internal argument is the theme:

(10) They fed peas to the baby. (Carrier and Randall 1992:193 (53b))

In this case, if adjectival passive formation applies, the output should be grammatical. However, the outputs of adjectival passive formation should also satisfy the theta-Criterion which requires all obligatory arguments to be realized. The derived adjectival structure **the fed peas* is ungrammatical because the second obligatory argument – *the baby* – remains unexpressed.

Levin and Rappaport (1986) note also a different condition illustrated by the verb *read* with two optional internal arguments: theme and goal. Whatever argument is externalized the output should be grammatical. However, as the examples below show, this is not the case:[4]

(11) a. The kids read the books.
b. the recently read books
c. *the recently read kids (Carrier and Randall 1992:194 (57a-b))

Levin and Rappaport (1986) state the following condition on the adjectival passive formation:

[4] The example (11a) is mine.

(12) Adjectival passive formation can apply to a verb only if it has a direct internal argument
(Carrier and Randall 1992:194 (58)).

Given this condition, Carrier and Randall (1992) use adjectival passive formation as a diagnostic for the argumenthood of postverbal DPs in resultative constructions. As the following examples show, adjectival passive constructions can be derived only from the resultative constructions based on transitive verbs:

(13) a. a wiped-clean table
b. *a drunk dry teapot (Levin and Rappaport Hovav 1995:43 (27))

### 2.3 Nominaliation

The rule of nominal formation applies to the resultative constructions based on transitive verbs but not to the resultative constructions based on intransitive verbs. Carrier and Randall (1992) focus on *–ing* process nominals because they retain the argument structure of the base verb (as opposed to result nominals). They assume that the rule of nominal formation, similar to the rule of verbal passive formation, suppresses the external argument of the base verb. The verbal passive and the nominal have the argument structure of their base verb, except that they do not have the external argument:

(14) a. Hungry students [$_V$ devour] vast quantities of junk food.
b. Vast quantities of junk food were [$_V$ devoured].
c. The [$_N$ devouring] of vast quantities of junk food takes no time at all.
(Carrier and Randall 1992:198 (63))

To determine the argument structure of process nominals Carrier and Randall (1992) analyze the DP immediately adjacent to the head – the *of-DP*. In the nominals derived from transitive verbs, this DP is interpreted as the direct internal argument of the nominal, inherited from its base verb (see (14) above). In contrast, in nominals derived from intransitive verbs this DP is an optional adjunct PP corresponding to the suppressed external argument:

(15) The rejoicing (of the villagers) lasted for days.
(Carrier and Randall 1992:199 (66))

The third interpretation of *of-DPs* corresponds to the *adverbial adjunct* interpretation. They are also PP adjuncts:

(16) The constant rejoicing of the holiday season makes Fred nervous.
(Carrier and Randall 1992:199 (67))

Carrier and Randall (1992) use the interpretation of *of-DPs* in nominals to determine the argument status of postverbal DPs in transitive and intransitive resultatives. Naturally, the *of-DP* of any nominal derived from a resultative construction can only be interpreted as a direct internal argument. When the nominal is derived from a transitive resultative this condition is satisfied, since the nominal preserves the direct internal argument of the base verb. But when the nominal is derived from an intransitive resultative this DP can only be interpreted as an adjunct since the base verb does not have a direct internal argument for the nominal to inherit. Therefore, nominals derived from intransitive resultatives are ungrammatical. The following examples illustrate this point:

(17) a. The gardener watered the tulips flat.
b. The watering of tulips flat is a criminal offense in Holland.
(Carrier and Randall 1992:173, 201 (1a) and (74a))

(18) a. The joggers ran their Nikes threadbare.
b. *The jogging craze has resulted in the running of a lot of pairs of Nikes threadbare.
(Carrier and Randall 1992:173, 201 (2a) and (74b))

The data outlined in this section show that middle constructions, adjectival passive constructions and nominalizations can be derived from resultative constructions based on transitive verbs but not from resultatives based on intransitive verbs. This suggests that the postverbal DPs in resultative constructions based on intransitive verbs are not arguments.

## 3.0 Russian prefix and the verb meaning

In this section, I present some data on Russian prefixation given in Townsend (1975) and Brecht (1985). They show that prefixation in Russian is a lexical process that has to do with meaning of the verb, its aspectual classification and argument structure. These issues are related as follows: prefixation in Russian often changes the meaning of the verb in unpredictable ways, that is, the meaning of the prefixed verb is not derived compositionally from the meaning of the prefix and the meaning of the stem. Meaning change may affect a verb's aspectual classification and argument structure. This, in turn, suggests the different argument status of the postverbal DPs in English resultative constructions and Russian prefixed verbs. I also discuss the difference between Russian prefixed verbs and English resultative constructions in terms of compositionality and productivity. I show that the meaning of an English resultative, which is a syntactically derived construction, is derived compositionally from the meaning of the main verb and the meaning of the resultative phrase and that the formation of resultative constructions, being a syntactic process, is more productive than the lexical process of prefixation.

### 3.1 Prefixation: meaning change

Townsend (1975:114-8) describes Russian verbal prefixes and the process of prefixation as follows. An unprefixed verbal stem is called a simplex stem. The majority of simplex stems in Russian are imperfective. With limited exceptions, the addition of a prefix perfectivizes a simplex imperfective stem. Traditionally, in most grammar books, Russian prefixes are regarded as either 'nonsemantic' or 'semantic'. The addition of a nonsemantic prefix to a simplex imperfective stem merely perfectivizes it and is regarded as the perfecitve partner of the imperfective verb:

(19) a. stroit' – PO-stroit'
build – buid.PERF
b. pisat' – NA-pisat'
write – write-PERF (Townsend 1975:116)

However, if the prefix is semantic, it not only perfictivizes the stem, but also changes its meaning. In this case, a new verb is not a perfective partner of the unprefixed stem and must build its own imperfective with an imperfectivizing suffix. This imperfective is sometimes called a 'second' or a 'secondary imperfective':

(20) a. stroit' – U-stroit' – U-stra-iva-t'
build – arrange – arrange.IMPF
b. pisat' – ZA-pisat' – ZA-pis-yva-t'
write – write down – write down-IMPF (Townsend 1975:117)[5]

[5] The examples of secondary imperfectives are mine.

Note that a given prefix may be nonsemantic for one stem but semantic for another:

(21) a. delat' – s-delat'
do-IMPF – S-do-PERF
'do – do'
b. prosit' – s-prosit'
ask-IMPF – S-ask-PERF
'ask (request) – ask (information)' (Townsend 1975:117)

However, Townsend (1975:117) points out that the number of prefixes which may serve as nonsemantic perfectivizers is very limited. The only prefixes which appear to have a purely perfectivizing function with any regularity are *po-* and *s-*.[6]

However debatable the issue of Russian prefixes is, Townsend (1975) maintains that it is still clear that prefixation in Russian is a process which involves some alteration in the meaning of the original simplex stem. He considers two types of prefixation: lexical and sublexical. A lexical prefix forms a new lexeme, usually related to one of the physical meanings of the original stem, or an abstract or qualitative meaning derived from it. In this case a secondary (derived) imperfective is built. In contrast, a sublexical prefix does not form a new lexeme but modifies the action in some way, usually with respect to time or intensity. The types of meaning involved here are usually called *Aktionsart* or 'mode of action'. Prefixed verbs formed by means of sublexical prefixes normally build no secondary (derived) imperfective. In general, the more the new perfective is felt to have independent meaning and not just aspectual or sublexical, the greater the chances of its having a derived imperfective.

Brecht (1985:14-6) also discusses the process of lexical and sublexical prefixation. He points out that atelic states and activities are normally represented in Russian by simplex stems which are mainly imprfective. It is possible to form different lexical items by adding various prefixes to these stems. This prefixation is a strictly lexical process and is accompanied by an automatic shift in the aspect of the verb. The new prefixed verbs represent different situations, specifically telic

[6] Townsend (1975:117) points out that some studies in fact questioned even whether the cases like *stroit' – po-stroit'* (build – build-PERF) and *delat' – s-delat'* (do – do-PERF) can be treated as true aspectual pairs, assigning to the prefix in these cases 'resultative' rather than purely perfective meaning (Isachenko 1962). Under this analysis, prefixation never results in mere perfectivization but always changes or modifies meaning of the original stem. Suffixal imperfictive derivation can be considered the only process producing true aspectual pairs.

ones – in the examples below, 'finding', 'rewriting', signing', etc. all imply a goal or endpoint. Brecht's (1985) example (14) is given here as (22):

| | | | |
|---|---|---|---|
| (22) a. | *pisat'* 'write' | PERE-pisat' | 'rewrite' |
| | | ZA-pisat' | 'jot down' |
| | | POD-pisat' | 'sign' |
| | | PRI-pisat' | 'ascribe' |
| | | VY-pisat' | 'copy out' |
| b. | *chitat'* 'read' | PERE-chitat' | 'reread' |
| | | ZA-chitat' | 'read out' |
| | | DO-chitat' | 'read up to' |
| | | VY-chitat' | 'find (in a book)" |

In contrast, a sublexical prefixation, or *Aktionsart* (mode of action) represents the focus on one component of a situation – its inception, conclusion, intensification, a limited period of its duration, etc. The sublexical prefixation also transforms the situation from atelic to telic by transforming an indefinite state or activity into achievement or culmination:

| | | | |
|---|---|---|---|
| (23) a. | *kurit'* 'smoke' | VY-kurit' | 'finish smoking' |
| | | DO-kurit'-sja | 'smoke to a climax' |
| | | ZA-kurit' | 'begin to smoke, light up' |
| | | ZA-kurit'-SJA | 'smoke too much' |
| | | NA-kurit'-SJA | 'smoke one's fill' |
| | | PO-kurit' | 'smoke for a while' |
| | | | (Brecht 1985: (15)) |

According to Brecht (1985), the crucial difference between lexical and sublexical prefixation is the character of the modification of the situation. Sublexical prefixation involves a semantic field which focuses on a component of the situation, whereas lexical prefixation creates an entirely different situation. As can be seen from this brief outline, the prefixation in Russian usually involves change in meaning and aspectual classification.

However, neither Townsend (1975) nor Brecht (1985) discusses explicitly a change of the argument structure of the verb as a result of prefixation. Table 1 below contains the verb *pisat'* 'write' combined with a number of different prefixes. These examples demonstrate that the arguments structure of a verb may change as a result of prefixation.[7] Note that all the prefixed verbs form their own imperfective:

Table 1: The Verb *pisat'* 'write'

| VERB STEM | PREFIX (PERFECTIVE) | SECONDARY IMPERFECTIVE | MEANING | OBJECTS |
|---|---|---|---|---|
| *pisat'* | - | - | write | letter, poem, book, paper… |
| | a. ZA-pisat' | ZA-pis-YVA-t' | 1. write down | answer, name, address… |
| | | | 2. record | concert… |
| | b. IS-pisat' | IS-pis-YVA-t' | 1.bring smth to its end by writing | 1.pen, pencil... |
| | | | 2. fill out with writing | 2.blackboard, notebook… |
| | c. POD-pisat' | POD-pis-YVA-t' | sign | agreement, document… |
| | d. O-pisat' | O-pis-YVA-t' | 1. describe | 1. person, situation… |
| | | | 2. circumscribe | 2.circumference, circle… |
| | c. S-pisat' | S-pis-YVA-t' | 1. copy | 1. text… |
| | | | 2. write off | 2. equipment… |

[7] Not all possible prefixes are included in the table and not all possible meanings for every combination are given. Also, the column 'Objects' does not contain a complete list of all possible arguments that may occur with the verb *pisat'* preceded by a particular prefix. The purpose of Table 1 is to show that prefixed verbs and thier base verbs alone may take different arguments.

As Table 1 shows, when the verb *pisat'* 'write' is used transitively it means 'to bring something into existence by writing'. This is a verb of creation and can take as its direct object anything which potentially can be brought into existence by the activity denoted by this verb: a letter, a poem, a book, etc. The prefix *iz-* changes the meaning of this verb in the sense that it is not a verb of creation any more. It is used to describe a different event. It means 'to bring something to its end by writing' and now it subcategorizes for a different kind of direct objects: this verb refers usually to surfaces that potentially can be used for writing or tools of writing – we have *is-pisat' ruchku/tetrad'* 'iz-write a pen/a notebook'. Spencer and Zaretskaya (1998) point out that when a prefixed verb takes a different direct object from its original stem, such an object has all the properties of a regular direct internal argument. The application of the tests for argumenthood, discussed in section 2 shows that this is indeed the case. Adjectival passive constructions, middle-type reflexive constructions and nominalizations can be derived from a prefixed verb *is-pisat'* 'iz-write' plus *ruchka* 'pen':

(24) a. is-pisannaja ruchka
IZ-written pen
'a pen that has run out of ink'
b. Takie ruchki ochen' bystro is-pisyvajut-sja
such pens very quickly IZ-write- REFL
'This sort of pen runs out of ink quickly.'
c. is-pisyvanie ruchek
IZ-writing pens-GEN
'the using up of pens' (Spencer and Zaretskaya 1998:24 (90d-e))

The examples in (24) suggest that in Russian, the direct object of a prefixed verb is an argument of the complex unit formed by the original stem and prefix.

### 3.2 Compositionality and productivity

I assume that the notions of compositionality and productivity are important for determining syntactic versus lexical nature of a construction. The meaning of a syntactic construction is derived from the meaning of its parts as opposed to a lexical construction, whose meaning often may not be derived compositionally. Besides, syntactic derivation tends to be more productive than lexical derivation. Given these assumptions, I shall now examine the difference between English resultative constructions and Russian prefixed verbs.

The crucial fact about semantics of English resultative constructions is that their meaning is derived compositionally, that is, it is predictable from the meaning of the main verb and resultative phrase. Though verbs impose various semantic restrictions on the resultative phrase that can appear with them – for example, as Spencer and Zaretskaya (1998) (12) observe, though we can say *wipe the glasses clean* we cannot say **wipe the glasses shiny* – under the assumption that we deal with the appropriate verb and adjective, in most cases, the meaning of the whole is derived compositionally. This is due to the fact that the main verb and resultative phrase are semantically specific. By definition, a resultative construction has to specify two components – a change of state and the activity which brought about this change of state. Thus, the meaning of the expression *she painted the door green* or *she wiped the table clean* is derived from the meaning of the verbs *paint* and *wipe* which define the activity, and the adjectives *green* and *clean* which specify the resulting change of state.

In contrast, in Russian, the meaning of the prefixed verb is not, in general, semantically compositional because Russian prefixes cannot be attributed a precise meaning on their own. According to Townsend (1975), most verbal prefixes have primary meanings of a physical, directional or spatial nature, often close to the meanings of the prepositions, to which they are historically related. Besides these primary meanings, many prefixes have one or several abstract meanings, whose connection with the primary sense may vary from obvious or remote to undistinguishable. Townsend (1975) also points out that many prefixed verbs are not worth analyzing in terms of possible meanings of the prefix because the prefix defies categorization entirely or because in order to categorize it we would have to come up with a prohibitively large number of meanings for a single prefix.

The notion of productivity is a tricky one. On the one hand, the process of prefixation in Russian may seem very productive. Any stem can be attached at least one prefix (usually there are more than two prefixes that can be combined with the same stem) and there is a large number of prefix-stem combinations. On the other hand, English resultative constructions may be claimed to be less productive. As already mentioned, verbs impose various restrictions on resultative phrases that can appear with them. However, although the number of prefix-stem combinations in Russian is indeed enormous, it does not seem possible to combine every stem with every prefix and there are no obvious reasons that would prevent certain stems from occurring with certain prefixes.

Meanwhile, the restrictions on resultative phrases that can appear with different verbs in resultative constructions in many cases can be explained. For example, Rappaport Hovav and Levin (1998) observe that result verbs cannot participate in resultative constructions:

(25) a. *Kelly broke the dishes off the table.
(meaning: Kelly broke the dishes and as a result they went off the table)
b. *Kelly broke the dishes into a pile.
(meaning: Kelly broke the dishes and made a pile out of them)
c. *Kelly broke the dishes valueless.
(meaning: Kelly broke the dishes and as a result they were valueless)
(Rappaport Hovav and Levin 1998:103 (9a), (10a) and (43a) respectively)

Rappaport Hovav and Levin (1998:105) propose that verb meaning is built up in a monotonic fashion, which precludes the elimination of any basic element of meaning and also the expansion of the meaning beyond the fully lexically-specified representation. They assume that manner verbs are mainly activity verbs which do not have a resulting state as a part of their lexical representation, whereas many result verbs are either achievements or accomplishments which have a resulting state as a part of their lexical representation. Resultative constructions involve the expansion of an activity to yield an accomplishment. The main verb normally denotes an activity (e.g. *wipe* or *paint*) and the addition of a resultative phrase extends the meaning of this verb to include a resulting state. Accomplishments, however, have fully lexically-specified representations. No additional expansion of their representation is possible since they represent the most complex representation available. Result verbs are normally accomplishments; accomplishments specify a change of state; no additional change of state can be introduced. Therefore, resultative constructions with result verbs such as *break* are ungrammatical.[8]

As we see, for both constructions there are restrictions on elements that can be combined. In Russian, it is impossible to combine every prefix with every stem to receive a valid output. Similarly, in English, not every verb and adjective can form a valid resultative construction. However, in Russian the combination of a prefix and a stem seems to be arbitrary whereas in English, it is possible to predict what kind of resultative phrase a verb is compatible with. In other words, if we confine ourselves to activity verbs and apply our knowledge of the world as to

[8] Levin and Rappaport Hovav (1995:55-62) discuss in depth various semantic restrictions on English resultative constructions.

what activities may potentially lead to what resulting states, the formation of English resultative constructions may be quite productive.

## 4.0 Prefixed verbs and verb-particle constructions

I argue that Russian prefixed verbs are much closer related to English verb-particle constructions than to English resultatives. As mentioned above, the similarity between prefixes and particles is the lack of a specific semantic content. English particles and Russian prefixes form a new semantic unit with the base verb, whose meaning often may not be derived compositionally. Just like Russian prefixed verbs discussed above, English verb-particle constructions often have a different meaning and different syntactic properties from their base verb.

### 4.1 Particles and prefixes

English particles, similar to Russian prefixes, often carry the directional meaning and correspond to the function of the homophonous prepositions:

(26) a. The soldiers ran up the hill and blew the bridge up.
b. Maria s-terla pyl' s polki.
Maria s-wiped dust-ACC from shelf
'Mary wiped dust off the shelf.'

Although English verb-particle constructions are claimed in some studies to be syntactic constructions (Neeleman and Weerman 1993, among others), they are more lexical in their behavior than English resultatives. Just like Russian prefixed verbs, the meaning of English verb-particle constructions often is not derived compositionally. Table 2 contains a set of verbs combined with different particles. Some particles seem to be more productive than the others; sometimes, the output is ungrammatical.[9] Note that though result verbs such as *break* do not participate in resultative constructions, they can form verb-particle constructions:

[9] The choice of verbs is arbitrary. Though native speakers' judgments vary regarding some of these examples, in general my informants were unanimous. I marked (!) those verb-particles constructions that were judged as ungrammatical or problematic by my informants but nonetheless were listed in the English Dictionary for Advanced Learners (Collins Cobuild, third edition 2001).

Table 2: English Verbs and Particles

| VERB | PARTICLE | | | | |
|---|---|---|---|---|---|
| | DOWN | IN | OFF | OUT | UP |
| a. break | break down | break in | break off | break out | break up |
| b. cry | *cry down | *cry in | cry off(!) | cry out | *cry up |
| c. grow | *grow down | grow in | *grow off | grow out | grow up |
| d. look | look down | look in | *look off | look out | look up |
| e. make | *make down | *make in | make off | make out | make up |
| f. run | run down | run in | run off | run out | run up |
| g. show | *show down | show in | show off | ?show out | show up |
| h. take | take down | take in | take off | take out | take up |
| j. wake | *wake down | *wake in | *wake off | *wake out | wake up |

English particles as well as Russian prefixes may affect a verb's aspectual classification and argument structure. Consider the examples below:[10]

(27) a. Sally wiped the table.
b. *Sally wiped.

(28) a. Sally wiped out on the ski hill.
b. *Sally wiped out the man/wiped the man out.[11]

---

[10] Thanks to Heather Bliss for these examples.

[11] The verb-particle construction *wipe out* can also mean 'destroy', in which case it would be transitive.

In (27), the verb *wipe* is obligatorily transitive as shown by ungrammaticality of (27b). Meanwhile, in (28) the verb-particle construction *wipe out* does not allow a direct object. Under the assumption that the verb in (28) is unaccusative, a change in argument structure is dramatic: with the addition of the particle, the verb changes from an obligatorily transitive verb whose external argument is an agent (at least in this case) to a verb which has only one argument and this argument is internal.

Another observation which suggests that English verb-particle constructions as well as Russian prefixed verbs are more lexical in nature than resultative constructions is that they can form an infinitive, while it is impossible with English resultative constructions – the latter cannot form conjoined verbs:

(29) a. I need to write off or fix my car.
b. You should wipe down or wash the table.

(30) Oni sobirajutsja po-chinit' ili s-pisat' staruju mashinu.
they are going PO-fix.INF or S-write.INF old.ACC car.ACC
'They are going to fix or write off the old car.'

(31) a. *She wants to paint green and sell her bike.
b. *She wants to wipe clean or wash the table.

In addition, English verb-particle constructions may be substituted by a morphologically simple verb:

(32) a. The soldiers blew the bridge up.
b. The soldiers exploded the bridge.

As the example below show, it is possible to replace some resultative constructions with a single verb as well:

(33) a. She hammered the metal flat.
b. She flattened the metal.

However, in this case, the replacement is not completely adequate, as the expression with a single verb refers only to the result and looses its activity component. In (33b), the metal became flat, but we do not know how she made it flat.[12]

The direct internal arguments of English verb-particle constructions also have properties of a regular direct object. Again, the application of the tests for argumenthood shows that middle constructions, adjectival-passive constructions and nominalizations are much better when derived from verb-particle constructions than from resultative constructions which involve transitive verbs:

(34) a. Wooden bridges blow up more easily than stone ones.
b. ?The cars of this type write off often.

(35) a. the blown-up bridge
b. ?/*the written-off cars

(36) a. The blowing up of bridges is prohibited.
b. ?The writing off of cars is a common thing.

Obviously, some of these examples are problematic. However, they are definitely better than the sentences derived from intransitive resultatives (see section 2 above). This suggests that English verb-particle constructions are intermediate between Russian prefixed verbs which are morphological words and resultative constructions which are syntactically derived complex predicates. Thus, middle constructions, adjectival passive constructions and nominalizations derived from English verb-particle constructions are not as bad as those derived from intransitive resultatives but they are still worse than those derived from Russian prefixed verbs.

Apparently, English verb-particle constructions, being lexically derived phrasal verbs, preserve some characteristics of their basic verb. The extent to which such sentences are problematic might depend on how far the meaning of the verb-particle construction is from the original verb and how conceptually different is the direct object of the verb-particle construction from the direct object required

---

[12] In some resultative constructions it is possible to substitute a verb + a resultative phrase by a single verb:

(i) a. She run herself ragged.
b. She exhausted herself.

However, in this case, the construction seems to be idiomatic in the sense that its meaning is not derived compositionally, since the verb *run* does not denote the actual running activity.

by the original verb. For example, *write off* does not preserve much of the meaning of its basic verb *write* and *car* is not a prototypical direct object of *write*. Thus, middle formation, adjectival passive formation and nominalization for *write a car off* are problematic. On the other hand, the verb-particle construction *blow up* seems to bear more conceptual similarity to its basic verb *blow* though *bridge* hardly could be considered as a prototypical direct object of this verb. It is not clear why in (34) – (36) the a-examples are better than the b-examples. However, here I am only interested in the contrast in grammaticality between middle constructions, adjectival passive constructions and nominalizations that are derived from intransitive resultatives and those derived from verb-particle constructions and I leave the account for this contrast for further research.

### 4.2 More about prefixes, particles and resultatives

Some additional observations suggest that neither Russian prefixed verbs nor English verb-particle constructions should be analyzed as counterparts of English resultatives. First, since the meaning of Russian prefixes and English particles is too vague they fail to specify a resulting state. Russian prefixes and English particles often transform an activity into an accomplishment or an achievement. However, as Levin and Rappaport Hovav (1995) point out, resultative constructions differ from simple accomplishments in that they lexically specify both the activity and the resulting state. In contrast, though Russian prefixed verbs and English verb-particle constructions are often accomplishments, they do not lexically specify the resulting state. For example, if we analyze the Russian expression *ona vy-terla stol* 'she *vy*-wiped the table' the same way as we analyze the English resultative construction *she wiped the table clean*, the Russian expression would mean that the table became *vy-* as a result of wiping. This is meaningless because the prefix *vy-* is too vague to define the change of state that the table undergoes. Similarly, the verb-particle construction *she wiped the table down* would mean that the table became down as a result of wiping. Let us have a look at the following examples:

(37) a. #Morris wiped the table clean but it is still dirty.
b. Masha vy-terla stol, no on ostalsja grjaznym.
Masha VY-wiped table-ACC but it remained dirty.
'Masha wiped the table but it remained dirty.'
c. Mary wiped the table down but it is still dirty.

The resultative construction in (37a) is semantically ill-formed whereas the Russian sentence with a prefixed verb in (37b) is not a contradiction. The English expression explicitly states that the table is clean now, but in Russian, this is only pragmatically presupposed. The Russian expression only says that she acted (presumably) to make the table clean and that she completed the action but it says nothing about how efficient her job was. The same is true about the English verb-particle construction in (37c). This sentence is not a contradiction for exactly the same reasons as the Russian sentence. Since it is impossible to attribute to the particle *down* any precise meaning on its own, we cannot define what exactly change of state the table has undergone.

Second, Russian prefixed verbs as well as English verb-particle constructions do not necessarily express a change of state. Compare the following sentences:

(38) a. Masha razo-grela sup.
Masha RAZ-warmed soup-ACC
'Masha warmed up the soup.'
b. Olga vy-myla chashki.
Olga VY-washed cups-ACC
'Olga washed the cups.'

(39) a. Mary rolled the carpet up.
b. Denis turned the page over.

(40) a. Claudia do-chitala knigu.
Claudia DO-read book-ACC
'Claudia finished reading the book.'
b. Morris pro-igral partiju v shaxmaty.
Morris PRO-played game in chess-ACC
'Morris lost a game of chess.'

(41) a. We looked the information up.
b. Tina thought the matter over.

In (38) and (39), the objects undergo a change of state. The soup, when it is warmed up, the cups, when they are washed, the carpet when it is rolled up and even the page when it is turned over, each undergoes a different kind of change. In contrast, the objects in (40) and (41) does not seem to undergo any obvious change of state. The book hardly undergoes any change of state when somebody finishes reading it and even less so a chess game when it is lost. Similarly, the information and the matter do not undergo any change of state when they are looked up and thought over respectively.

These observations put together suggest that Russian prefixed verbs and English verb-particle constructions do not work exactly as English resultatives. They should not be given the same analysis because they do not have the two essential characteristics of the resultative construction. First, prefixes and particles, which presumably correspond to the resultative phrase, cannot precisely describe the result because their meaning is too vague. Second, since the meaning of prefix verbs and verb-particle constructions is not derived compositionally and may often be different from that of the base verb, direct objects of Russian prefixed verb constructions and verb-particle constructions does not necessarily undergo a change of state.

## 5.0 Conclusion

The crucial assumption of this paper is that the syntactic behavior of a verb is determined by its meaning. Given this assumption, English resultative constructions on the one hand, and English verb-particle constructions and Russian prefixed verbs on the other hand, have different syntactic properties. As follows from the analysis above, the latter two share a very important semantic feature. English particles and Russian prefixes cannot be attributed any precise meaning on their own; they comprise a single semantic unit with the verbs they are combined with. This new lexical unit often has a different meaning from that of the original verb and, as a result, may exhibit different syntactic properties. In contrast, in English resultative constructions, the meaning of the two components of a complex predicate – a verb and a resultative phrase – is precisely defined. The meaning of the verb within the English resultative construction does not change with an addition of the resultative phrase and selects the same arguments as in isolation. Therefore, postverbal DPs within English resultative constructions on the one hand and English verb-particle constructions and Russian prefixed verbs on the other hand have different argument status. In the case of English resultative constructions based on intransitive verbs they are not arguments but unselected objects. However, in Russian, they are arguments of prefixed verbs and exhibit all the properties of normal direct objects. In English verb-particle constructions, the argument status of such DPs is less obvious. Some examples of middle constructions, adjectival-passive constructions and nominalization are problematic. This is perhaps due to the fact that English verb-particle constructions are phrasal verbs and constitute a kind of intermediate category between English syntactic resultatives and Russian morphologically derived prefixed verbs. Besides, as suggested in section 4, English verb-particle constructions and Russian prefixed verbs cannot be treated as counterparts of English resultative constructions because they do not specify a resulting and their direct objects do not always undergo a change of state.

**Acknowledgements**
I thank Malka Rappaport Hovav for her comments on the earlier version of this paper and Betsy Ritter for generously spending her summer time reading numerous drafts of this paper and her invaluable comments and suggestions. I am also grateful to my classmates, who provided me with their judgments and examples.

**References**


Brecht, Richard D. 1985. 'The form and function of aspect in Russian'. In Michael S.Flier and Richard D.Brecht (eds). *Issues in Russian Morphosyntax*. Ohio: Slavica, Columbus. 9-33.

Burzio, Luigi. 1986. *Italian Syntax*. Reidel: Dordrecht.

Carrier, Jill and Janet H. Randall. 1992. 'The argument structure and syntactic structure of resultatives'. *Linguistic Inquiry* 23, 173-234.

Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, Massachusetts: MIT Press.

Isachenko, Aleksander V. 1962. *Die Russische Sprache der Gegenwart*. Halle: T.I.Formenlehre.

Keyser, Samuel Jay and Thomas Roeper. 1984. 'On the middle and ergative constructions in English'. *Linguistic Inquiry* 15, 381-416.

Levin, Beth and Malka Rappaport. 1986. 'The formation of adjectival passives'. *Linguistic Inquiry* 17, 623-61.

Levin, Beth and Malka Rappaport Hovav. 1995. *Unaccusativity*. Cambridge, Massachusetts: MIT Press.

Neeleman, Ad and Fred Weerman. 1993. 'The balance between syntax and morphology: Dutch particles and resultatives'. *Natural Language and Linguistic Theory* 11, 433-75.

Rappaport Hovav, Malka and Beth Levin. 1998. 'Building verb meanings'. In Miriam Butt and Wilhelm Geuder (eds). *The Projection of Arguments*. Stanford: Center for the Study of Language and Information. 97-134.

Rappaport Hovav, Malka and Beth Levin. 2001. An event structure account of English resultatives. Language 77 (4), 766-97.

Simpson, Jane. 1983. 'Resultatives'. In Lori Levin, Malka Rappaport and Annie Zaenen (eds). *Papers in Lexical-Functioanl Grammar*. Bloomington: Indiana University Linguistics Club. 143-58.

Spencer, Andrew and Marina Zaretskaya. 1998. 'Verb prefixation in Russian as lexical subordination'. *Linguistics* 36, 1-39.

Townsend, Charles E. 1975. *Russian Word-Formation*. Cambridge, Massachusetts: Slavica Publishers Inc.

Williams, Edwin. 1981. 'Argument structure and morphology'. *The Linguistic Review* 1, 81-114.

Zalianjak, Anna and Aleksej Shmelev. 1997. *Lekcii po Russkoj Aspektologii* (Lectures on Russian Aspectology). München: Verlag Otto Sagner.

Ilana Mezhevich
Department of Linguistics
University of Calgary
2500 University Drive NW,
Calgary, Alberta
T2N 1N4 Canada
imezhevi@ucalgary.ca

## The Acquisition of English Onsets: The Case of Amahl[1]

Rebecca Hanson
University of Calgary

### Abstract

The acquisition of English onsets by one English-learning child is examined in close detail, with particular focus on the acquisition of /s/ and /s/-clusters. The observation that target /s/ in harmony environments is sensitive to the feature [labial] as opposed to [coronal] and [dorsal] provides support for a feature geometry model in which [labial] versus [lingual] is a possible distinction, e.g. Brown (1997). Further, the unique behavior of target /s/ in the developing phonology motivates the proposal that physiological factors, such as articulatory difficulty, can have consequences in the grammatical system. In particular, it is proposed that a constraint against lingual continuants, which require a precise physical coordination that may not have yet developed, can account for the patterns in the child's acquisition of /s/ clusters. A comparison with the acquisition of /f/, a non-lingual continuant, and that of /l/, another lingual continuant, provides further support for this proposal. The conclusions reached here are consistent with notion of a phonological system grounded in independent, functional principles as argued for in, for example, Goad (1997).

[1] **Notational conventions:** IPA symbols are used for transcriptions. However, I follow Smith's (1973) bracketing convention: slanted brackets '/ /' are used for phonemes of the target language (English); straight brackets '| |' represent phonemes of A's developing system; and square brackets '[ ]' represent phonetic realizations. I also use Smith's subscript dot diacritic (superscript on 'g') to indicate a 'voiceless lenis' articulation: i.e. [ḅ, ḍ, ġ].

## 0. Introduction

This thesis is primarily concerned with the acquisition of onsets, both singleton and cluster, as observed in one child learning English. The particular focus is on fricatives, with an attempt to account for the difference between the treatment of target /f/ and target /s/ in both onset types. The thesis begins with a presentation of the theoretical backdrop for the rest of the paper. Chapter 1 offers an overview of first language acquisition as well as introducing some of the commonly observed processes found in child language. In Chapter 2, I turn to general phonological theory. I outline the relevant aspects of feature geometry theory, present the specific geometry model I adopt, that of Brown (1997), and motivate that choice. This is followed by a return to acquisition, focusing on the feature-geometric approach to phonological development.

Chapter 3 marks the beginning of the case study itself. There I provide a summary of the first seven stages of Amahl's acquisition of singleton onsets according to the data in Appendix C of Smith (1973). Facing some unaccountable patterns in the treatment of /s/, I propose a phonetically-motivated constraint against lingual continuants and suggest that early articulatory difficulties with the pronunciation of /s/ trigger a repair strategy which gradually decreases in linguistic significance. That is, in the stages considered here, the first effect of the constraint is to delete the /s/ altogether; next, only the offending feature, [continuant], delinks; later, the /s/ is finally produced but it is phonetically longer than the target. The hypothesis of an articulatory constraint proves useful in accounting for the behavior of /s/ both in singleton and in cluster onsets.

Amahl's acquisition of onset clusters in the first seven stages is addressed in Chapter 4; here again the proposed constraint helps to account for the unique treatment of target /s/-clusters.

In Chapter 5, I present some data from later stages and demonstrate how they provide support for the constraint against lingual continuants. I show how the effects of this constraint weaken, phonologically speaking, as acquisition proceeds. Some possible implications of this sort of constraint are addressed in Chapter 6.

## 1. Overview of Acquisition

There has been a steady increase in the interest in acquisition data for phonological theory, especially with the more widespread acceptance of the assumption of continuity between child and adult grammars. The continuity hypothesis claims that although child phonologies are constantly changing, developing toward the target grammar, each change in the developing language conforms to universal principles; at no point in the acquisition process will a child diverge from these principles and use rules or constraints that never occur in adult languages. It must be the case, therefore, that any rule, process, representation, or constraint proposed for a child's grammar is also found in adult language (though not necessarily the target). This corresponds to what Macken (1995) calls the "strong identity hypothesis," which claims that the same fundamental capacity underlies child and adult speech, and language acquisition involves "enhancing" this capacity rather than learning it from scratch.

I likewise assume continuity in this thesis. In addition, I assume that throughout the acquisition process, input from target grammar directs the acquisition path within the confines set by Universal Grammar (UG). Specifically, each elaboration of the child's grammar brings it closer to the target grammar at some level. With respect to the phonology, Ingram (1996) has formalized this latter assumption at the segmental level in his Distinctive Feature Hypothesis (given in (1)), offering a straightforward check for avoiding incompatibility or discontinuity between the developing and developed grammars:

(1) The Distinctive Feature Hypothesis

> Any feature assigned to a child's representation must be present in the target phonemes.

For example, Mike (from Pollock, 1983) produces [n] for target /n, w/ and [d] for target /d, s, ð/. There are two ways that the [n] vs. [d] contrast could be represented: i) with the feature [sonorant] – [n] is [+sonorant] and [d] is [–sonorant]; ii) with [nasal] – [n] is [+nasal], [d] is [-nasal]. According to the Distinctive Feature Hypothesis, however, only (i) is compatible with the input; (ii) does not capture the fact that [n] also substitutes for the [-nasal] target, /w/. On the other hand, [sonorant] captures the [n]/[d] contrast and at the same time is consistent with all the target phonemes: [n] is [+sonorant] like /n, w/, and [d] is [-sonorant] like /d, s, ð/ (Ingram, 1996). Only by referring to the Distinctive Feature Hypothesis can continuity be maintained in ambiguous situations like these.

### 1.1 Phonological processes in acquisition

It is well known in the study of first language acquisition that there is a significant amount of uniformity: that is, certain phonological processes are consistently observed both between languages and within a given language. These can be generally summarized under two categories: syllable-related processes (prosodic level), and segment-related processes (melodic level). The former have been described as driven by a UG-supplied, bimoraic Minimal Word template (Fee, 1995; Salidis & Johnson, 1997). The latter involve feature markedness issues and universal phonological processes. The following sections address each of these levels separately before considering their interaction.

### 1.2 Syllable-level processes

Fee (1995, 1996) and Demuth & Fee (1995, cited in Demuth, 1996) propose the following stages of prosodic development:

| | | |
|---|---|---|
| (2) | Stage I | Core (CV) Syllables |
| | Stage II | Minimal Words ($\sigma_\mu\sigma_\mu$ or $\sigma_{\mu\mu}$) |
| | | a. core syllables – (C)VCV |
| | | b. closed syllables – (C)VC |
| | | c. vowel length distinctions – (C)VV |
| | Stage III | Stress-Feet |
| | Stage IV | Prosodic Words |

In Stage I, children produce what Fee (1996) calls "subminimal" words: monomoraic CV utterances, which tend to show variable length and tension in the vowel. At this stage, commonly-observed processes include Coda Deletion (3a) and Syllable Deletion (usually of the unstressed syllable; see (3b) for an example).

(3) a. coda deletion: 'dog' → [da]

b. syllable deletion: 'away' → [we]

At Stage II, the Minimal Word template requires the presence of two morae in every word. Very often, partial and total reduplication of a single CV syllable will be observed at this point (4a and 4b, respectively). Another common repair strategy, especially in Stage IIa, is epenthesis; for example, in (4c) a schwa is inserted to make the target word into a CVCV form.

(4) a. partial reduplication: 'Peter' → [bibə] (Fee, 1995)
b. total reduplication: 'bottle' → [baba]; 'father' → [fafa]
c. epenthesis: 'pig' → [pigə]; 'blue' → [bəlu]

By Stage IIb, coda segments are no longer deleted; however, consonant clusters are simplified, generally in favor of the least sonorous element (Fikkert, 1994) as in (5) below. The relative sonority of segments, or sonority hierarchy, will be presented in chapter 2 (sections 2.4 and 2.5); for the examples in (5) we can simply note that /s/ is more sonorous than /t/ and /l/ is more sonorous than /k/. Thus, /t/ and /l/ survive their respective cluster reductions.

(5) 'stop' →[tɑp]; 'milk' → [mik]

### 1.3 Segment-level Processes

On this level, the representation of distinctive features plays the key role. A more detailed discussion of the internal structure of segments can be found in chapter 2.1; here, I will only refer briefly to a few of the main points. First, I assume that segments do have internal organization, and are composed of features drawn from a finite, UG-supplied set. Second, these features are organized in a hierarchy which reflects both markedness (less hierarchical structure = less marked) and phonetic dependency relationships. Finally, as I discuss in chapter 2.3, I assume that UG supplies a minimal structure which is then elaborated in response to cues in the input. For first language acquisition, these assumptions account for the general observations summarized in (6), taken from Macken (1995:676):

(6) In general, the unmarked consonants:

a) are acquired first
b) are most frequent in the child's lexicon
c) have the fewest restrictions on their distribution
d) serve as replacements for the corresponding marked consonants during the stage when the contrast is neutralized.

The unmarked segments have the least hierarchical structure, therefore (6a); typologically, less marked segments are more common in the world's languages, likewise (6b); finally, neutralization of a contrast goes in favor of the less marked, thus (6d). In general, stops are less marked than fricatives and coronals less than non-coronals; acquisition processes associated with these have been called Stopping and Fronting, respectively (Ingram 1976, cited in Fee 1995):

(7) a. Stopping: 'juice' → [dut]; 'this' → [dɪ]
b. Fronting: 'go' → [do]

However, the situation is not as simple as the generalizations in (6) would make it appear. As acquisition progresses, phonological processes such as feature harmony, or assimilation, become more significant. For example, while [d] substituted for [g] in (7b), we find [d] assimilating to [g] in (8). These kinds of examples have often been used to support theories of feature underspecification both in acquisition and in adult languages.

(8) 'desk' → [gɛk] (Smith, 1973)

Assimilation is quite widespread in child language; the examples in (9) show how it can be adjacent or non-adjacent, progressive or regressive:

(9) a. Adjacent assimilation
progressive: 'bump' → [bumb]
regressive: 'sweet' → [fweet]
b. Non-adjacent assimilation
progressive: 'doggie' → [gagi]
regressive: 'noisy' → [nojni]

In these cases, there could be more than phonetic assimilation at work. Possibly, at this stage of development there is a restriction at some level against having more than one Place or Manner feature specified. This restriction could trigger assimilation as a repair strategy, but it does not itself select which segment must assimilate; that, presumably, is chosen according to universal rules of feature spreading (see chapter 2.2)

**1.4 Interaction between syllabic and segmental levels**

Much attention has recently been paid to the ways in which the syllabic and segmental levels interact in acquisition. In a very general sense, it can be said that they do interact, and that prosodic structure does influence segmental elaboration. Slobin (1973) observed that some prosodic positions were more salient to children, and summarized his observation in the phrase, "children pay attention to the edges of words". The interplay between syllable and segment acquisition is illustrated quite clearly in metathesis (the reversal of sounds within a word), another process often observed in child language. Examples of metathesis are supplied in (10).

(10) a. 'spaghetti' → [pəsgɛDi]
b. 'animal' → [amənal]

In (10a), a consonant cluster which violates the Sonority Sequencing Principle (see chapter 2.4) is dealt with through metathesis (thus correcting the violation) as well as epenthesis (reducing the onset members to one). Alternatively, (10a) could be seen as the result of a feature alignment constraint which motivates the movement of [labial] to the left edge of the word. This latter proposal can also account for the metathesis in (10b). Such feature alignment patterns have been noted by several researchers. For example, Fee (1995) notes that segments are first pronounced target-like in syllable-initial position; Macken (1996) in a similar vein discusses how marked features emerge first (and most accurately) word initially. Velleman (1996) deals specifically with metathesis in an OT framework, attributing it to the ranking of feature alignment constraints relative to constraints involving faithfulness to the input. Dinnsen (1996) argues that feature markedness depends on syllable position, and demonstrates compatibility and continuity between child and target language using Kiparsky's (1993) context-sensitive radical underspecification.

Since evidence exists for the context-sensitivity of features, in my analysis of A's inventory I have restricted my attention to a single syllable position: in order to avoid the complications of context effects, only word-initial onsets were considered. This decision is partly motivated by the assumption that a feature can be contrastive in one position but not in another. Evidence that this is in fact the case can be found from pairs like those in (15), which show that at the same stage in Amahl's acquisition (stage 8), [voice] was contrastive in codas (15a) but not in onsets (15b).

| | | | | |
|---|---|---|---|---|
| (11) | a. | [lɛg̊] 'leg' | vs. | [luk] 'look' |
| | b. | [d̥ɔiː] 'toy' | vs. | [d̥aun] 'down' |

To summarize: this chapter has presented some of the commonly-observed ways in which child language deals with difficult aspects of the target language. Given the theoretical importance of the continuity hypothesis, an effort has been made to relate the child and target languages in terms of repair strategies such as deletion, epenthesis and assimilation. In the next chapter, I turn to the relevant areas of linguistic theory which aim at accounting for observations like the ones summarized thus far.

## 2. Theoretical Background

In this chapter, I summarize the theoretical assumptions which form the backbone of my analysis of the data in the following chapters.

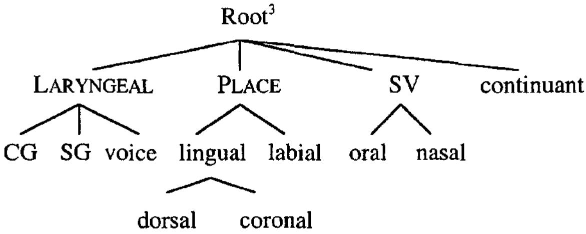
### 2.1 Internal structure of the segment[2]

Following a common approach in phonological theory, I begin with the assumption that phonological features are not unordered bundles, as in SPE, but are organized in a UG-defined structure known as the feature geometry. While several different models have been proposed, the formal properties of the geometry are generally constant. The particular model I will assume is adapted

[2] Some of the material in this chapter has appeared in an earlier form in Hanson (to appear)

according to Brown's (1997) revisions of the Rice & Avery (1995) geometry. In Section 3 below I will address this model more closely. Note, however, that I do depart from Brown (1997) in placing [voice] as an explicit dependent of the Laryngeal node; this decision is based on acquisition data, and will be motivated in detail in chapter 5.

(12)




As first proposed by Clements (1985), and modified in Rice & Avery (1991), there are two major node types beneath the Root node, which gathers together the segment as a unit. Organizing nodes, indicated by small caps in (12), generally correspond to the anatomy of the vocal tract, and organize features into natural classes or constituents. They dominate the content nodes, which have actual articulatory status such as 'Labial' or 'Spread Glottis', and which in turn may dominate secondary content nodes or terminal features. I adopt the stance that these features are monovalent, with contrasts represented through the presence or absence of a node rather than through binary '+' and '-' values. I further assume that redundant or predictable features, as determined by universal markedness patterns, are absent from underlying representations – that is, they are underspecified – and may be filled in at the level of phonetic implementation by a default fill-in rule.

Within the geometry, nodes are hierarchically related to each other in a dependency relationship. In (12), for example, [spread glottis] is a dependent of the Laryngeal node; this reflects the fact that the activity of the feature [spread glottis] entails the activity of the Laryngeal node. The choice for encoding a particular relationship as one of dependency can be motivated by phonetic

[3] CG = [constricted glottis]; SG = [spread glottis]; SV = Sonorant Voice

considerations (as with the previous example) or by phonological patterns, which are often best captured through underspecification. For example, typologically, the presence of labials and velars in a system implies the presence of coronals; this pattern is reflected in most geometry models by designating [coronal] as the default interpretation of the Place node. In other words, in the unmarked case, coronals are represented with a bare Place node and the feature [coronal] is not specified until phonetic implementation (for further discussion see Paradis & Prunet (eds), 1991).

Finally, the feature geometry is universal, able to capture all existing phonological systems, though no one language will exploit the entire structure.

### 2.2 Phonological processes

An important advantage of the feature geometry over linear models is the ability it has to capture a wide range of phonological processes with a limit set of well-defined operations. These operations are commonly gathered under three types: Spreading, Delinking, and Fusion (Avery & Rice 1989).

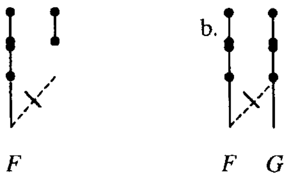
The process of Spreading is schematized in (13) and its general conditions, as summarized by Avery & Rice, are given in (14).

(13)


(14) a. Spreading can occur only if a structural target is present
b. A feature or node can spread only to an empty position

Adopting this view of Spreading (others exist; see for example Steriade 1987) has important implications for my approach to A's data. For example, (14a) rules out any analysis that relies on node generation, such as the spreading of [oral] to a segment that lacks an SV node. (14b) further limits Spreading to a feature filling role rather than a feature changing one ; in other words, it cannot trigger the delinking of a feature that is a dependent of the same node. (15) below illustrates some situations where Spreading fails according to (14): in (15a), the feature (F) fails to spread because the structural target is absent; in (15b) it is because the target position is already filled by another feature, G.

(15) a. b.




"Delinking" refers to the operation that breaks an association line between a feature or node and its superordinate node. It is commonly used to account for instances of neutralization such as word-final devoicing, which involves delinking the Laryngeal dependents in a specific environment (Avery & Rice 1989).

Finally, Fusion is an OCP-driven operation, reducing two identical adjacent elements to one. Avery & Rice define fusion as "an operation which takes identical primary content nodes and fuses them provided that the nodes are non-distinct; i.e. both do not dominate different secondary nodes." Fusion, or coalescence, will not be important in the analysis presented below; however, it has been an important process in other analyses of onset cluster reductions. For example, Hanson (1999) proposes a fusion analysis of A's acquisition of /s/-clusters; Chin & Dinnsen (1992) offer a two-step version of feature coalescence, which they analyze as involving first Spreading and then deletion or degemination; and Gnanadesikan (1995) looks at coalescence in onset clusters from an Optimality Theoretic perspective.

## 2.3 Acquisition of the feature geometry

If the feature geometry is, as I assume, a reasonable model of phonological representation, any proposed hierarchy must also be able to account for language acquisition. It is particularly important that a model do so without violating the principles of learnability and continuity that are important to all areas of child language theory.

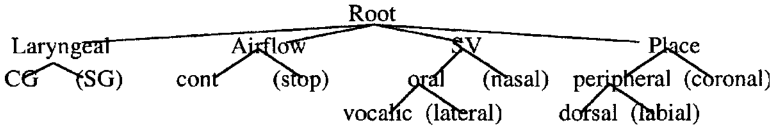
There are two main theoretical approaches to the acquisition of the feature geometry, which Brown (1997) refers to as the "Pruning" and "Building" hypotheses. The Pruning hypothesis assumes that UG supplies a fully differentiated structure like the one in (12) above, capable of representing all possible phonological contrasts. As the child realizes that not all of these possible contrasts are used in the ambient language, the unnecessary structure is "pruned" away. A closely related proposal is what Rice & Avery (1995) call the "full specification" approach to segmental acquisition, which claims that the segments themselves are acquired as units, fully specified for its particular universal features. Underspecification can still be achieved, however, by later pruning away the features that are redundant in that particular system. A disadvantage of the pruning approach is that it relies heavily on negative evidence: adjustments are made to the developing geometry only when it becomes clear that a given contrast is not relevant, or that a given feature is redundant, in the target grammar.

According to the Building Hypothesis, on the other hand, UG supplies only a minimal structure, which is elaborated in response to positive evidence of a relevant contrast in the input. In addition, the geometry is expanded in a step-by-step fashion, or monotonically.

Like Brown and Rice & Avery, I adopt the Structure Building approach; in addition to its reliance on positive evidence, it best captures both the uniformity and the variability so often observed in language acquisition. The uniformity is a result of the deterministic aspect of the geometry when approached with a structure-building view: a higher node (e.g. an organizing node) must be acquired before any of its dependents (e.g. content nodes and/or features). Variability appears due to the individual child's freedom to begin acquisition at any organizing node, and to continue by elaborating either within or outside that node. In other words, the acquisition process is constrained in the general order, but free in the specifics.

Recently, acquisition studies have begun to play an important role in evaluating proposed feature geometry models. As I mentioned in section 2.1 above, the geometry I will adopt here was originally proposed by Rice & Avery (1995) specifically to account for child data, and was revised by Brown (1997) on the basis of further acquisition research. The proposed revisions were to the SV and Place node. For comparison with (12), the Rice and Avery geometry is supplied in (16).

(16) Rice & Avery (1995)




cont = [continuant]; CG = [constricted glottis]; SG = [spread glottis]; SV = Sonorant Voice; brackets indicate default interpretations

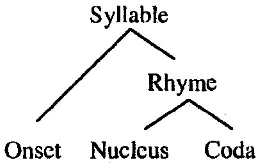
The most obvious difference between (12) and (16) is the lack of the Airflow node in (12). Since the status of this node, proposed by Rice and Avery, is still uncertain and does not figure crucially in this work, I have left it out of the adopted model (12). The revisions of the SV node are not important in this thesis and will not be addressed here except in noting that they were motivated by acquisition data.

The revisions to the Place node, however, will be especially relevant in chapter 3 below. A comparison of (12) and (16) indicates that Brown has dramatically revised Rice & Avery. Like for the SV node, this revision was motivated by acquisition data: many children acquired a labial/velar contrast before acquiring the coronal/non-coronal contrast, an observation which cannot be predicted from the Place node structure in (14). Therefore, Brown departed from Rice & Avery, proposing that the lingual articulations, [coronal] and [dorsal], should be grouped together as opposed to [labial], similar to independently motivated articulation-based geometry models such as Browman & Goldstein's (1986, 1989).

### 2.4 Internal structure of the syllable

In line with current phonological theory, I assume that the syllable is a phonological constituent, and that it has internal structure. As with the feature geometry, there are several different models of syllable-internal structure (see Blevins, 1995, for a summary); in this paper I will adopt the model presented in (17), though the details of the structure are not crucial.

(17)




Each of the lowest subcomponents – onset, nucleus and coda – can optionally branch, allowing for consonant clusters (branching onset or coda) or for long vowels and diphthongs (branching nucleus).

The syllable has been referred to as a phonological domain containing exactly one sonority peak (Blevins 1995). The importance of sonority in constraining the possible sequences of segments within a syllable has been repeatedly confirmed in phonological research. While it is not altogether clear what the phonetic correlate of sonority is (beyond a general reference to acoustic saliency), in phonological terms it is possible to rank segments according to their sonority. A typical sonority hierarchy for English phonemes is given in (18).

(18) sonority scale (adapted from Carr 1993)

vowels > j, w > r > l > m, n > ð, v, z, ʒ > θ, f, s, ʃ > b, d, g > p, t, k

That is, vowels are the most sonorant segments, and sonority decreases towards the voiceless stops, which have the lowest sonority value. Of particular relevance here is the role of the Sonority Sequencing Principle (SSP, also called the Sonority Sequencing Generalization) in determining what is and is not a possible onset cluster. A version of the SSP is provided in (19), taken from Blevins (1995).

(19) Sonority Sequencing Principle

> Between any member of a syllable and the syllable peak, a sonority rise or plateau must occur.

In other words, sonority rises towards the nucleus of the syllable, with the preference being that the edges of the syllable will have the lowest sonority possible. However, as we will see in Chapter 4, sonority isn't always respected. In particular, /s/ is notorious for violating the SSP; for example, /s/-stop clusters (e.g. /sp, st, sk/), the more sonorant /s/ is further from the nucleus than the stop and yet these clusters are very common in the world's languages. The problem of /s/ in clusters has been extremely persistent; it continues to be a theoretical puzzle and has been dealt with in various ways. Clements & Keyser (1983) propose that /s/ is extrasyllabic and thus is not part of the onset at all; a similar view is taken by Kaye et al. (1990), who claim that /s/ is in the coda of a null syllable preceding the other 'cluster' members. Phonetically, /s/ could be granted "quasi-syllabic" status because it compares to typical syllable nuclei in its acoustic salience and intrinsic length (Dobrovolsky, p.c.). This quasi-syllabic behavior shows up, for example, in English paralinguistic utterances like 'psst'. In this thesis, I propose that the precision necessary for its articulation can also contribute to the unusual behavior of /s/.

## 2.5 Acquisition of syllable structure

In an extensive study of the acquisition of syllable structure in Dutch, Fikkert (1994) found that the SSP played a central role in onset cluster reduction, and that acquisition seemed to follow a sonority-directed path. The various acquisition paths she recorded can be resolved into two main strategies, given in (20).

(20) a. Maximize sonority distance between onset and rhyme
b. Maximize sonority distance between members of the onset cluster

(20a) captures the tendency for the least sonorant member of the target onset to survive cluster reduction (cf. (5) in chapter 1.2); (20b) captures the observation that when onset clusters are produced by the child, there is a tendency for the elements of the cluster to be as far apart in sonority value as possible even if that causes a divergence from the target.

Up to this point, we have looked at observational patterns of L1 acquisition (chapter 1) and presented relevant aspects of linguistic theory (this chapter). Beginning with the next chapter, the remainder of this thesis will be devoted to a single case study, looking in detail at one child's early stages of onset acquisition.

## 3. The Acquisition of Singleton Onsets

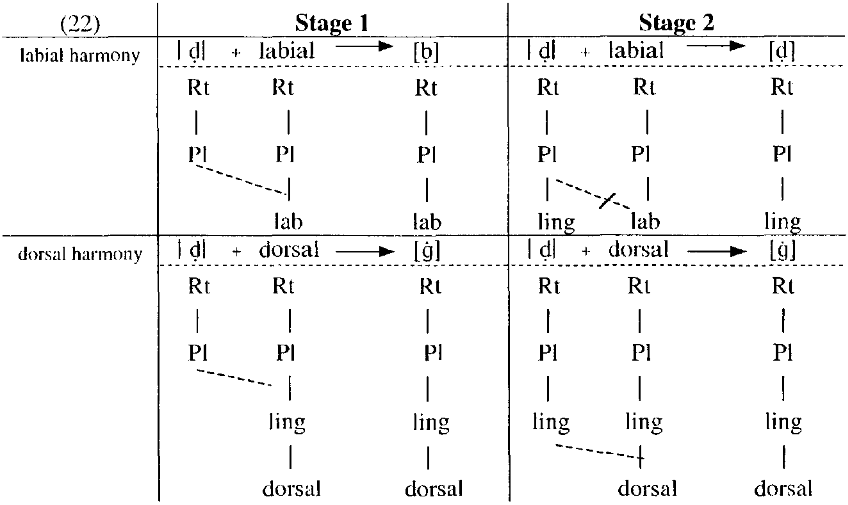
This chapter looks in detail at the early steps taken by one child in the acquisition of singleton onsets. The data is taken from Smith's (1973) diary study of Amahl (henceforth 'A'), with a focus on the first seven stages as delineated by Smith.

### 3.1 Summary of the data

When Smith's study begins, A has already made some progress towards the target geometry. A survey of his productive inventory through stages 1-7 is provided in (21) along with the target phonemes that each segment replaces; on the far right is the feature which, according to the Distinctive Feature Hypothesis (see (1) in Chapter 1 above), is assumed to be specified on the given segment.

| (21) | **Amahl** | **Target(s)** | **Feature** |
|---|---|---|---|
| a. | \| b̥\| | /p, b/ - also /t/, stage 1 only | [labial] |
| b. | \| d̥\| | /t, d, s, z, l/ | ([lingual]) |
| c. | \| g̊\| | /k, g/ - also /s/ and /l/ under harmony | [velar] |
| d. | \| w\| | /f, w/ - also /s/ and /l/ under harmony | [cont]; [labial] |
| e. | \| m\| | /m/ | [nasal]; [labial] |
| f. | \| n\| | /n/ | [nasal] |

To explain the parentheses in (21b): in Stage 1, there is one instance of target /t/ being realized as [b̥] (e.g. [b̥ebu] 'table'), though by Stage 2 /t/ resists labial harmony while still being subject to velar harmony (e.g. [g̊e:k] 'take'). I assume, then, that at Stage 1 the coronal stop was represented with a bare Place node, thus being vulnerable to the spreading of any Place features. At Stage 2, however, [lingual] is specified; from that point on spreading is blocked in the case of [labial] but not [velar], which is also a dependent of [lingual]. This development of the Place node is illustrated in (22).

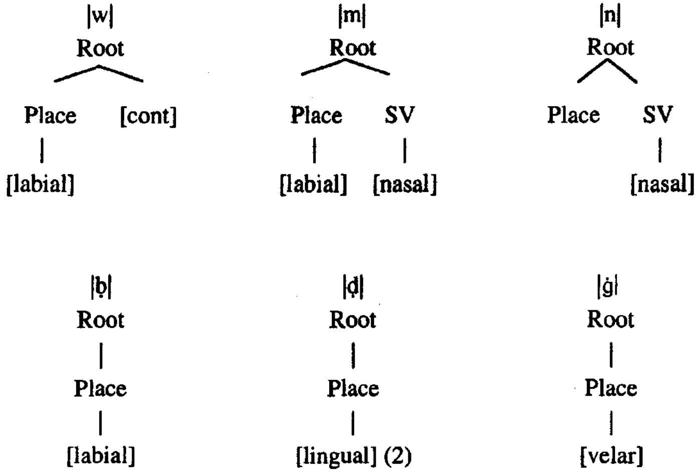



Using the Rice and Avery model (see chapter 2.3.2), it is not clear how to capture harmony in which [velar] but not [labial] can spread to an underspecified coronal.

From the information in (21), we can see that A has elaborated the Place node to distinguish the three major places of articulation. The SV node, with its [nasal] dependent, is also contrastive. Note, however, that A's |w| cannot be specified [SV], since target /f/, which it substitutes for, is not [SV]. I assume, therefore, that |w| has been analyzed as a labial continuant which is phonetically realized by A as an approximant. Obstruent voicing is not yet contrastive.

The underlying representations of each of (21a) – (21f) are provided in (23) below to illustrate the parts of the geometry A is working with at this point. In (23b), the feature [lingual] is absent at Stage 1 but present from Stage 2 onward. This same feature is deliberately absent from the coronal nasal (23f) out of preference for underspecification: even without an elaborated Place node, |n| is distinct from the rest of A's inventory.

(23)




With this information about the correspondence between A's phonemes and the target ones, we can now look more closely at his phonological system. Substitution patterns in particular can provide valuable insights into this system; the next few sections therefore will focus on A's substitutions for target /s/ and /l/.

**3.2 The split realization of /s/**

Target /s/ shows an interesting pattern of replacement by A. We saw in the previous section (in (21) above) that /s/ can be realized predictably as [ḍ], [ġ] or [w]. The examples in (24) illustrate.

(24) a. 'see' [ḍi:] 'side' [ḍaiḍ]
b. 'sock' [ġɔk]
c. 'soft' [wɔpt] 'same' [we:m]

In (24a) and (24b), we see /s/ realized as a stop, coronal in isolation or in a coronal environment, and velar in a velar environment. So far, it seems like this is straightforward consonant harmony, with /s/ underlyingly no different from the coronal stop in terms of both manner and place features. This hypothesis is strengthened when we look at the realizations of /t/ (25) in coronal and velar environments: they are identical with the realizations of /s/.

(25) a. 'tie' [d̥ai] 'tight' [d̥ait]
b. 'talk' [g̊ɔk]

However, when we look at the labial environment, there is an unexpected divergence. In (24c) above, we saw that target /s/ is realized as [w] if [labial] is specified in the same word. With the target /t/, on the other hand, we find the expected labial stop [b̥] as the output (26).

(26) 'table' [b̥e:bu]

Furthermore, recall from section 3.1 above that the coronal stop assimilates to the labial only in this one instance in all of Smith's database – and only for the first stage. By Stage 2, labial harmony no longer applies to the stops. For /s/, however, labial harmony continues until Stage 10. Clearly, then, /s/ and /t/ are different for A, as illustrated in the following summary of their respective substitution patterns (note that these are not intended as formal rules; the format is used for convenience only):

| | | | |
|---|---|---|---|
| (27) | /s/ → [w]/ __ (...) [labial] | /t/ → [b̥] / __ (...) [labial] (Stage 1 only) |
| | → [g̊] / __ (...) [velar] | → [g̊] / __ (...) [velar] |
| | → [d̥] elsewhere | → [d̥] elsewhere |

Although they are only distinct in one environment, the fact that /s/ and /t/ are subject to different substitutions indicates that they must have different underlying representations. However, it is surprisingly hard to pinpoint exactly what the underlying difference is: since /s/ is realized as [d̥] in non-harmony conditions, it seems that it must have a UR identical with that of [d̥]. Likewise, the fact that both emerge as the velar stop [g̊] under velar harmony confirms this hypothesis. Unfortunately, /d/ becomes [b̥] under labial harmony, while /s/ becomes [w] – so perhaps /s/ is underlyingly [continuant], distinguished from [w] only by place features. Why, then, should it be realized as [d̥] in isolation?

There are two features to which we could reasonably attribute the divergence between /s/ and /t/ under conditions of labial harmony. First, the feature [labial] itself is obviously important. The other relevant feature is [continuant], since it distinguishes the two output forms, [w] and [b̥]. In order to better understand how these two features contribute to A's treatment of /s/, it would be helpful to try to account for his choice of substitutions; in the next section, one possible account is brought forward and evaluated.

### 3.3 A Constraint against lingual continuants

The following proposed account for the realization of /s/ makes crucial use of Brown's (1997) revision of the Place node. Recall from chapter 2.3.2 that in this model, [coronal] and [velar] are subsumed under the node [lingual], which along with [labial] is a direct dependent of the Place node. Both |d̥| and |g̊| are specified for [lingual], while |b̥| is not (rather, [labial]).

I propose that the split substitution of /s/ can be accounted for by positing a low-level, articulatory constraint against lingual continuants; under this proposal, lingual continuants are targeted for repair for articulatory reasons. Thus, the proposed constraint has a phonetic motivation: it is easier to articulate a labial approximant continuant than a coronal (or dorsal) fricative continuant, since fricatives require greater precision in their articulation (Ladefoged & Maddieson, 1996:329). The consequent repair could come in the form of delinking [continuant], as discussed immediately below, or by deleting the /s/ altogether – as, for example, in the stage 1 productions of 'soap' and 'soup':

(28) 'soap' [u:p] 'soup' [u:p]

The strategy of deleting /s/ will appear again in A's reduction of /s/-clusters (chapter 4); for now, let's look more closely at the implications of the proposed constraint.

Under this hypothesis, |s| is in fact specified as [continuant] underlyingly. However, it can only surface as a continuant if it is non-lingual (i.e. [labial]) – that is, under conditions of labial harmony – since at the level of phonetic implementation, any representation specified for both [continuant] and [lingual] will be repaired by delinking [continuant]. This is illustrated in (29), where the asterisk on the left indicates that the structure is (temporarily) disallowed in A's system.

(29)

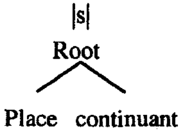
| *Root | | Root |
|---|---|---|
| Place continuant | *Repair*: delink [continuant] → | \| |
| \| | | Place |
| lingual | | \| |
| ... | | lingual |
| | | ... |

It is reasonable to assume that [continuant] should delink rather than, say, the Place node, for two reasons. First, A's system already includes lingual non-continuants, so the repaired structure is consistent with his phonology; this would not be true if Place delinked since /h/ (the Placeless continuant) has not been acquired and /l/ has its own difficulties (see 3.4 below). Second, it is [continuant] that is causing the phonetic difficulty: A has no difficulty producing stops at all places of articulation. Thus, delinking [continuant] is the simplest way to repair the offending representation. It is important, however, to note that delinking is not triggered for labial continuants, which do not require such fine motor control to articulate.

With this foundation laid, the split replacement of /s/ can be accounted for in the following way. Recall that /s/ becomes [w] only if [labial] is present within the same word. This substitution, at first glance, appears to be easy enough to account for under the current hypothesis: it is simply due to feature spreading. Since the result of spreading yields a geometry containing [labial] and [continuant], (29) does not apply and the target /s/ is realized as a labial continuant: [w]. However, this solution raises another problem: how can [labial] spread to the position already occupied by [lingual]? Only a feature-changing operation could accomplish this, but in a theory of monovalent features as we are assuming here it is preferable to limit Spreading only to a feature-filling role (Brown, 1997). Thus, an account which avoids feature changing is to be preferred over one that depends on it.

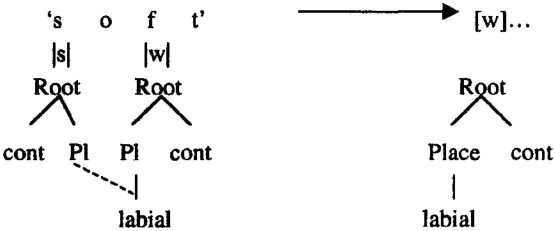
In this case, it is relatively simple to avoid the feature-changing operation, if we assume that in A's system |s| is represented as in (30), with a bare Place node (compare the proposal for |n| in (23) above):

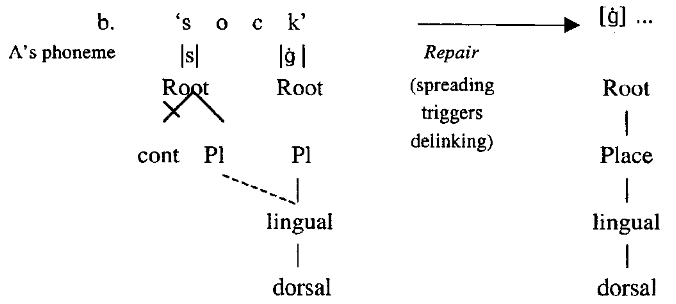
(30)

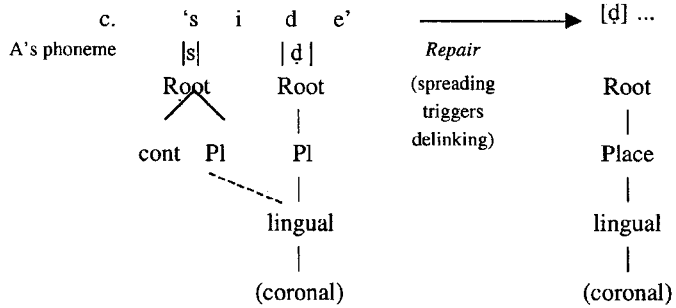



Further Place specifications are received either through Spreading (of [lingual] plus dependents, or of [labial]) or through a default fill-in rule providing [lingual] and implemented as [coronal]. If Spreading provides [labial], |s| is realized as [w], as in (31a). If Spreading provides [lingual]-[dorsal], the structure qualifies for repair and [continuant] delinks, yielding the dorsal stop [ġ] as in (31b). Finally, if Spreading (or the default rule) provides [lingual]-[coronal], again the repair strategy applies, yielding the coronal stop [ḍ] as in (31c). Note that in this hypothesis we are committed to order the default rule before the repair strategy.

(31) a.
A's phoneme






We will find some further support for the proposed constraint against lingual continuants in chapter 4, from the behavior of /s/ in cluster reduction strategies. Now, let's look briefly at A's treatment of /l/, which is quite similar to that of /s/.

### 3.4 The realization of /l/

The situation of target /l/ is even more difficult to pinpoint than that of /s/, especially because of the widespread disagreement in the literature regarding the representation of laterals. Levin (1988), Rice & Avery (1991), Piggott (1992, 1993), Yip (1990) and Brown (1997) all propose different models to account for various patterns in the crosslinguistic behavior of lateral segments. Brown (1997) provides a survey of the existing versions, and proposes that the feature [lateral] is not a phonological primitive; instead, laterals emerge phonetically from a certain combination of features (i.e. a bare Place node plus either the SV node (for sonorant [l]) or the Airflow node (non-sonorant [ɬ])). Because these feature combinations can easily be produced by spreading operations, in acquisition data it is difficult to discern whether a surface lateral is contrastive or simply the result of consonant harmony. As we will see in a moment, A's acquisition data do not make the problem of laterals any easier to solve. Let's begin, then, by simply noting the different realizations of target /l/ in these early stages; these patterns will prove important in chapter 4.

Table 1 organizes some representative examples of A's substitutions for target /l/ in different feature environments. Unfortunately, there are no words available to illustrate /l/ in isolation; 'lady' and 'lazy' come the closest with only coronals in the word. 'Lamp' and 'left' illustrate the realization of /l/ in the presence of [labial], while 'lick' and 'leg' illustrate the [velar] environment. Blank cells in the table indicate stages where no data is available and '>' indicates that the onset has reached consistent target production.

Table 1: A's production for target /l/ in various environments

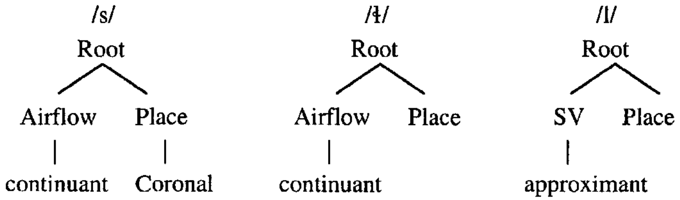
| | Stages | | | | | | |
|---|---|---|---|---|---|---|---|
| Example | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| '/l/ady' | ḍ | ḍ/l | | l > | | | |
| '/l/azy' | ḍ | | | l > | | | |
| '/l/amp' | | | | | w | | |
| '/l/eft' | | | w | | | | |
| '/l/ick' | | | | ġ | | | ġ/l |
| '/l/eg' | ġ | | | | | | |

From this table we can see that there is a striking similarity between A's treatment of /l/ and that of /s/. Both harmonize to place of articulation, with the same output: [w] under labial harmony, [ġ] under velar, and [ḍ] in the presence of coronals.

Using the Optimality Theory framework, Goad (1997) addresses in detail A's treatment of /l/ in these harmony environments. To account for the patterns, Goad offers the following proposal: while [labial] and [approximant] can be licensed on a single segment [coronal] (or [dorsal]) and [approximant] cannot. Taking into account the functional equivalence between the features [continuant] and [approximant] (Brown 1997), this proposal is practically identical to the constraint proposed here, which permits labial but not coronal and dorsal continuants.

The similarities between /s/ and /l/ in A's system are intriguing; unfortunately, a more detailed look at them would go beyond the scope of this thesis. However, further research might find an advantage in using Brown's model, in which the representations of /s/ and the laterals are very similar:

(32)




Because of his nearly identical treatment of target /s/ and /l/, the patterns in A's data support the approach in (32) more than one in which [lateral] is a feature dependent on the SV node and /l/ has little or nothing in common with /s/.

## 4. The Acquisition of Onset Clusters

Given the proposals in Chapter 3 for A's geometry elaboration in these early stages, I now turn my attention to his acquisition of onset clusters. Again, the data is taken from stages 1 to 7 in Smith (1973). Since target complex onsets are not acquired until stage 9, we can expect to find A employing consistent repair strategies to deal with this deficit. We can also expect to gain more information regarding the underlying representations of segments in A's inventory by looking at how different segments interact with each other and with the available syllable structure.

### 4.1 A's onset simplification

Before looking in detail at these early stages in A's acquisition of onset clusters, I want to draw attention to the fact that stop- and fricative-initial clusters are not treated the same way. The division between the two types of obstruent-initial cluster is illustrated in (33) comparing 'play' versus 'fly' and 'slide'.

(33) a. 'play' [ḅei]
b. 'fly' [wai] 'slide' [lait]

The stop-initial cluster /pl/ is reduced to a plosive, while the fricative-initial /fl, sl/ is apparently reduced in favor of the sonorant. This latter reduction is especially interesting, since it goes against the preferred sonority sequencing. As discussed in chapter 1.2 above, acquisition is expected to follow a sonority-directed path; that is, onset cluster reductions should be in favor of the least sonorous element in order to maximize the sonority rise between onset and nucleus. From (33) we see that the stop-initial clusters conform to this generalization, but the fricative-initial clusters do not. Thus, in the following sections I will address each of these separately.

## 4.2 Stop-initial clusters

A's stop-initial clusters follow straightforwardly along the predictions of sonority, and so do not require much discussion. Representative data for all the target stop-initial clusters are provided in Table 2. Note that in each case it is the plosive that survives the reduction while the more sonorant liquid is deleted. From (33b) above we know that this pattern is not due to an inability to produce onset liquids and can best be attributed to the SSP.

Table 2. Reduction of target stop-initial clusters

| Example | Target cluster | A's production |
|---|---|---|
| 'play', 'black' | /pl/ or /bl/ | [ḅ] |
| 'precious', 'bread' | /pr/ or /br/ | [ḅ] |
| 'tray', 'drive' | /tr/ or /dr/ | [ḍ] |
| 'clean', 'glue' | /kl/ or /gl/ | [ġ] |
| 'cloth', 'green' | /kr/ or /gr/ | [ġ] |

The lack of voicing contrast that we saw above (chapter 3.1) in the singleton onset stops appears here in the clusters as well. It is reasonable to assume, then, that A repairs the target clusters simply by deleting the sonorant; the remaining stop is, as expected, specified only for place of articulation.

### 4.3 Fricative-initial clusters

The fricative-initial clusters are not so simple to deal with: not only do they seem to reduce contrary to sonority expectations, in this section we will see that not even all fricative clusters behave the same way.

It turns out that the /f/-clusters are the easiest to address, so we will start with them. Although their substitute, the highly sonorant [w], is unexpected, at least it is always the same. As Table 3 illustrates, regardless of the second member of the cluster, and regardless of the featural environment, /f/-clusters are always reduced to [w].

Table 3. Reduction of target /f/-initial clusters (Stages 1-7)

| Example | A's production |
|---|---|
| '/fl/ag' | [w] |
| '/fl/ower' | [w] |
| '/fr/uit' | [w] |
| '/fr/og' | [w] |

Recall that singleton /f/ was also realized as [w]; it looks, then, like /f/ clusters are treated just like the stop-clusters we saw in 2.1 above. That is, they are reduced in favor of the less sonorant element (/f/), which is then treated exactly as if it were a singleton onset. The [w] cannot be the realization of the liquid, since /l/ is realized as [w] only in the presence of a labial (/r/ behaves the same way in this matter). From examples like 'flag' and 'fruit', then, we can conclude that it is the /f/, not the liquid, which survives the reduction: the liquid would have been realized as [g̊] in 'flag' and as [d̥] in 'fruit'.

There is an alternative analysis, which allows that it is the liquid that survives. Since /f/ is labial, it is possible that the feature [labial] spread from /f/ to the liquid, yielding [w] as with singleton /l/. Subsequent deletion of the /f/ would leave just [w] in the onset as observed. However, this analysis is less favored for two reasons. First, it is more complicated than the previous analysis, involving a rule ordering of consonant harmony before deletion rather than simply deletion;

and second, it contradicts the SSP while the previous analysis followed the predictions of sonority. It is more reasonable to assume that the liquid was deleted due to a preference for a sharper sonority rise from onset to nucleus; the fact that the actual output, [w], does not conform to sonority preferences is simply a consequence of A's phonological system at this point.

Now, let's turn to /s/-clusters. Singleton /s/ onsets, we saw, acted differently from both stops and /f/ - or rather, acted like both, depending on the environment. In terms of cluster reduction, however, stop- and /f/-initial clusters behaved the same way; does /s/ conform to the pattern, or does it diverge again?

As we will soon see, /s/ is once again treated differently than we might expect. Since /s/, like /f/ and the stops, is assumed to be lower in sonority than liquids, nasals and glides, we would expect to find that /s/ is consistently the survivor of reduction for /s/-sonorant clusters. For /s/-stop clusters, on the other hand, we would expect the stop to survive. Unfortunately for sonority considerations, this is not the actual pattern.

Table 4 below summarizes A's output for each target /s/-clusters (only two-member clusters are addressed).

Table 4. Reduction of target /s/-initial clusters (Stages 1-7)

| Target cluster | Example | A's production |
|---|---|---|
| /sp/ | 'spit' | [b̥] |
| /st/ | 'stay' | [d̥] |
| | 'stuck' | [g̊] |
| | 'stop' | [d̥] / [b̥] (stage 1 only) |
| /sk/ | 'sky' | [g̊] |
| /sl/ | 'sleep' | [w] |
| | 'slide' | [l] |
| | 'slug' | [l] |
| /sm/ | 'small' | [m] |
| /sn/ | 'snail' | [n] |
| | 'snake' | [ŋ] |

Before discussing the information in Table 4, it would be useful to recall the substitution patterns of each relevant segment when it is alone in the onset. These patterns are summarized in (34).

(34) a. /s/ → [w]/ ___ (...) [labial]
→ [ġ] / ___ (...) [velar]
→ [ḍ] elsewhere

b. /p/ → [ḅ]

c. /t/ → [ḅ] / ___ (...) [labial] (Stage 1 only)
→ [ġ] / ___ (...) [velar]
→ [ḍ] elsewhere

d. /k/ → [ġ]

e. /l/ → [w]/ ___ (...) [labial] (simplified)
→ [ġ] / ___ (...) [velar]
→ [ḍ] elsewhere

f. /m/ → [m]

g. /n/ → [n]

Now we are in a better position to evaluate A's reduction strategy. First, consider the /s/-stop clusters in Table 2. These are realized as unvoiced stops, which is not surprising since both /s/ and stops are usually produced that way; but we do have evidence that it is in fact the stop which survives cluster reduction. The crucial example is 'stop', where /st/ is realized as [ḍ] (alternating with [ḅ] at stage 1) in the environment of [labial]; this is how /t/, but not /s/, acts in a singleton onset. Assuming that A employs the same strategy for all /s/-stop clusters, we can thus conclude that these clusters are reduced in favor of the stop. So far, sonority predictions are borne out just as with stop- and /f/-initial clusters.

The surprises show up in the /s/-sonorant clusters. Turning now to target /sl/, once again it appears that we will not be able to tell which of the two cluster members remains, since /s/ and /l/ are so much alike in their behavior at this point. With 'sleep', for example, it is unclear how to determine which one of the members is ultimately being realized as [w]; both are equally likely candidates. Again, however, there is more explicit information available from other examples.

'Slide', which is produced with an [l] onset, doesn't prove very helpful. At the point when this production of 'slide' was recorded, both /s/ and /l/ are known to surface as [l]:

| | | |
|---|---|---|
| (35) | 'slide' → [lait/d̥ait] | 'sit' → [lit] |
| | 'light' → [lait] | 'little' → [d̥idi:/lidi:] |

However, since it was more usual for /l/ to become [l] than for /s/, we can tentatively propose that it is /l/ which survives the reduction – though more support is certainly necessary.

That support comes from considering A's production for target 'slug'. This word was first recorded at stage 10, unfortunately outside the current range of stages, but still useful – particularly when we take into account the realizations of target /s/ and /l/ by that point. At stage 10, /s/ is no longer ever realized as [l], but /l/ always is. Thus, since /sl/ug surfaces as [l]ug, we can assume that the /s/ is deleted and the more sonorant /l/ is retained. Again, assuming that A's phonological system is consistent, we can hypothesize that all /sl/ onsets have been repaired in this way.

This hypothesis gains strength when we look at /s/-nasal clusters. From Table 4 we know that these clusters surface as a nasal; thus, there is sufficient evidence to propose that all /s/-sonorant clusters are reduced in favor of the sonorant. This is, according to sonority, a dramatic difference from the treatment of all other onset clusters and deserves a closer look.

### 4.4 /s/-clusters and the constraint against lingual continuants

From the /s/-clusters it is apparent that sonority sequencing is not the only factor in determining which member of the onset will survive reduction: regardless of whether the output maximizes the sonority curve, /s/ is deleted. However, we saw with the other onset clusters that the SSP was the deciding factor. The fact that /f/-clusters obeyed the SSP while /s/-clusters did not is further support for the hypothesis that /s/ itself is problematic for A (rather than, say, fricatives in general). Thus, if we assume that there is a phonetically-motivated constraint against lingual continuants, the unique treatment of /s/-clusters falls out naturally. It is no surprise that these clusters should be reduced by deleting the problematic /s/, even if the result is less favored according to sonority. And since the constraint does not apply to labial continuants, /f/-clusters, like stop-clusters, are free to reduce according to sonority preferences.

In summary, then, a consideration of A's treatment of target onset clusters has revealed that in all but one case, clusters are reduced in favor of the less sonorant element. The one exception is in the target /s/-clusters, which are always reduced by deleting the /s/ regardless of the relative sonority of the surviving element. This 'deletion' repair strategy is consistent with the constraint against lingual continuants which was motivated by the treatment of target /s/ in singleton onsets.

## 5. A Glance at Later Stages

The previous two chapters focussed on the first seven stages of A's acquisition of onsets. We saw that the realization of /s/ as either a stop or as [w], depending on the environment, posed some problems if we wished to maintain continuity assumptions along the lines of the Distinctive Feature Hypothesis. In response to these difficulties, I proposed am articulatory constraint that targets lingual continuants for repair. In this chapter, I will present some developments beyond stage 7 which provide further support for this proposal.

### 5.1 Development of the obstruent voicing contrast

The elaboration of the Laryngeal node occurs in the following way. First, at Stage 11, obstruent voicing begins to appear on target voiced stops (looking at word-initial onsets only):

(36) 'bath' [ba:t] 'bottle' [bɔġəl] (cf. [ḅɔkəl] at stage 10)

Target voiceless stops, however, generally remain as they were before – voiceless lenis, or perhaps better, 'unvoiced' in contrast with the voiced stops.

(37) 'pink' [b̥ig̊] 'pig' [b̥ig]

Throughout the next stage, the non-voiced stops gradually become phonetically voiceless, often alternating between non-voiced and voiceless for a while according to Smith's transcription; but it is not until stage 13 that they become aspirated in the appropriate context. The appearance of aspiration indicates that the feature [spread glottis] ([SG]) has been acquired.

The development of the voiceless half of the target voicing contrast is illustrated in (38) with 'pink'.

| | | | |
|---|---|---|---|
| (38) | 'pink' | [b̥ig̊] | (Stage 11) |
| | | [b̥iŋk/piŋk] | (12) |
| | | [pʰiŋk] | (13) |

These observations have implications for the structure of the Laryngeal node. First, modal voicing emerged at a stage where target voiceless stops were still produced as they were before voicing became contrastive. This indicates that the voiced segments receive the new feature specification, and that it is therefore necessary for [voice] to be in the geometry as a dependent of the Laryngeal node. The voicing contrast at stage 11, then, would be represented as in (39).

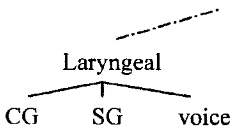
(39)

| /p, t, k/ | /b, d, g/ |
|---|---|
| \|b̥, d̥, g̊\| | \|b, d, g\| |
| Root | Root |
| \| | \| |
| Laryngeal | Laryngeal |
| | \| |
| | voice |

This kind of representation neatly captures the continuity between the previous stages (when [voice] was not specified) and the target voiceless stops in stage 11 – both of which have the same phonetic manifestation. Based on this argument, then, I would propose a revision to the Laryngeal node in the Rice & Avery and Brown geometries, placing [voice] explicitly as a dependent to represent modal voicing. This revision, illustrated in (40) for Rice & Avery, differs from most long-standing models of the geometry (for example: McCarthy

1988, Clements & Hume 1995) only in that I assume all features to be unary rather than binary.

(40)




At stage 11, then, only the feature [voice] is specified; the unstable realization of the voiceless stops at stage 12 (see (38)) I attribute to an attempt to make the distinction between the voiced and unvoiced versions more acoustically distinct. At stage 13, the addition of the feature [spread glottis] allows aspiration to appear – and also has a significant effect on the voiceless fricatives, which I will address in the next section.

**5.2 Development of fricatives**

Stage 13 marks an important point in A's development of voiceless fricatives; now, /f/ begins to appear as target in onset position:

(41) 'full' [wul/vul/ful] 'first' [fəːt] (stage 14)

Something very different happens with target /s/, however, as the examples in (42) illustrate. The previous replacement pattern for /s/, discussed in chapter 3.1, suddenly changes, and /s/ is now realized as either an aspirated stop or an affricate.

(42) 'say' [t$^{s}$ei] 'same' [t$^{h}$eim] 'sock' [k$^{h}$ɔk/t$^{h}$ɔk]

Comparing (41) and (42), it becomes apparent that once again the labial and coronal continuants are being treated differently: /s/ always is realized as either an aspirate or an affricate, and /f/ never is – it never surfaces as [p$^{h}$] or [p$^{f}$], for example. I will return to this issue shortly, but first I want to address the fact that aspiration and voiceless fricatives (at least, /f/) are acquired at the same time.

It is very common in feature grids to see voiceless fricatives specified [-voice], or lacking [voice], except for [h] which is universally specified [SG] like the aspirated stops . This kind of model, however, cannot capture the simultaneous acquisition of voiceless fricatives and aspiration at stage 13. A's acquisition data strongly suggest that voiceless fricatives are also specified [SG]. Independent support for this specification can also be found in both phonological and phonetic research.

The phonological support comes from a variety of sources. For example, Clements (1985) explicitly places [SG] in his representation of /s/. Lévesque (1992) argues that the specification of [SG] on fricatives greatly simplifies the autosegmental account of the behavior of coda-/s/ in Seville Spanish. Hanson (2000) proposes that this same specification helps to unify aspiration and sonorant devoicing patterns in English onset clusters. Finally, the common process of debuccalization, where /s/ becomes [h], can be easily accounted for under this hypothesis through the delinking of Place features (McCarthy 1988).

The phonetic evidence for [SG] on voiceless fricatives is even more striking. Laryngeoscopic studies show that the glottis is in fact open wide during the production of [s] (Dobrovolsky, p.c.). Hirose (1999) likewise notes that there is a tendency for the glottis to open further for [s] than for stops.

Furthermore, there are articulatory advantages to having a spread glottis during the production of fricatives, which involve a precise – tight but not closed – stricture in the vocal tract. Opening wide the glottis allows more air through, making it easier to maintain a supraglottal air pressure high enough to keep air moving through the resistant, narrow opening.

Therefore, based on A's acquisition and supported by these phonetic and phonological observations, I will assume that voiceless fricatives are specified for the feature [SG] .

But even if this is the case, how does it help us account for the latest developments in the problematic /s/? In particular, it does not explain why there is such a long delay between when /f/ reaches target and when /s/ does. I propose that this particular divergence is, like the others we have seen, attributable to the articulatory constraint against lingual continuants, and that the addition of [SG] to the distinctive feature inventory has an important impact on how the constraint is implemented. In the next section, the arguments for this proposal are presented.

### 5.3 Spread Glottis and the constraint against lingual continuants

We noted above that increasing the volume of high-pressure air through the oral cavity greatly helps in the production of frication. It is feasible that this kind of assistance could weaken the effect of the articulatory constraint proposed here. One phonetic obstacle would have been removed, namely the difficulty in producing a turbulent airstream with the limited amount of air that a 'voiceless lenis' vocal fold setting allows into the oral cavity. Thus, once [SG] is specified underlyingly the constraint against lingual continuants can relax to the point that it is no longer obligatory to delink [continuant].

One advantage of this approach is that it provides a phonetic motivation for the sudden change in a phonetically-motivated constraint. Another is that it allows that the constraint itself is only weakened, not removed. This point is important in order to account for the fact that it is another ten stages before /s/ is produced as target. If the constraint was done away with at this stage, why the delay? On the other hand, if the constraint is not affected by the acquisition of [SG], why the sudden change in the realization of /s/ so soon afterwards? It is important to recall that /f/ has begun to reach target production, indicating that the difficulty is, once again, with the lingual continuant alone.

But again, if this is the case, then why is /s/ realized as a stop – aspirated or affricated, true, but a stop nonetheless? Since [continuant] is no longer delinked, why isn't it at least a continuant of some sort? This question is especially relevant when we see that the voiced coronal fricative is also affricated at stage 13:

(43) 'zoo' [d$^z$u:]

In order to get a better understanding of what is happening with /s/, and so to answer these questions, it is particularly relevant to note that target affricates are not realized as affricates at this point. The examples in (44) show how they are realized as stops well past stage 13 (the numbers in parentheses indicate stages).

| | | | | | | |
|---|---|---|---|---|---|---|
| (44) | 'chair' | [tɛː] | (13) | 'giant' | [dæːnt] | (19) |
| | | [tˢɛə] | (21) | | [dzaint] | (28) |

Arguably, then, the "affricates" produced by A in place of /s/ are not true affricates; there is no reason to believe that they are phonologically complex segments. This observation, I believe, suggests that A's [tʰ] and [tˢ] can reasonably be considered continuants. That is, I propose that the affrication and aspiration portions actually represent A's primary attempt at the lingual continuant, with the [t] portion being an excrescent stop inserted to make the articulation easier – similar to glottal stop epenthesis in vowel-initial words but with a purely phonetic purpose.

Under this hypothesis, there is no need to claim that the stop portion of the articulation has any phonological status. Rather, we can keep the representation for /s/ that we have used so far, and assume that it continues to be underlyingly specified for [continuant]. In this way continuity can be maintained.

To summarize, in this section I have argued that the acquisition of [spread glottis] allows a relaxation of the constraint against lingual continuants. With this constraint weakened, the feature [continuant] is able to influence the articulation of target /s/ to a greater extent than it could before. As a result, instead of a bare stop replacing /s/, at this stage the replacement is a stop-initiated continuant: either [tʰ] and [tˢ], in free variation.

Before we move on past this stage, recall that the target voiced fricative /z/ also appeared as an 'affricate' at stage 13 (see (43) above). This does not cause any difficulty for the current account: if the constraint is weakened, then all lingual continuants should show the effects, not just the [SG] ones.

### 5.4 Gradually relaxing the constraint

The analysis so far implies that the removal of the proposed articulatory constraint is a gradual process. If that is the case, then there should ideally be more evidence of its relaxation before /s/ finally reaches target production. In fact, there are two pieces of data from later stages that provide such evidence.

First, and perhaps most striking, are the following stages in the development of 'see':

(45) 'see' [tʰi:/tˢi:] (13)
[ᵗsi:/si:] (21)
[si:] (22)

Note especially stage 21, where Smith transcribes the stop portion with a superscript, indicating that it is now less robust than it was earlier. This kind of phonetic weakening of the stop is just what we would predict if the [tˢ] of previous stages was a "stop-initiated continuant" as I claimed in the preceding section. The fact that the next step after [ᵗs] is the target production is also expected.

Another support for the gradual relaxation of the constraint against lingual continuants comes from the two observed instances of lengthening the fricative in target /s/-clusters at stage 23:

(46) 'Smith' [m̥əuk/məuk/s:m̥əuk]
'snail' [n̥eil/s:n̥eil]

Since the lengthening in these cases alternates with the absence of the /s/, it is reasonable to assume that it is simply a consequence of the difficulty in pronouncing that segment. In other words, the proposed constraint has by this stage been weakened to the point that it could hardly be called a constraint – more like a cautionary notice, a flag to indicate that this segment requires more attention than others.

The examples in (46) draw attention to another puzzling aspect of A's onset clusters: the production of voiceless sonorants in the place of target /s/ clusters. This is a consistent strategy between stages 15 and 24. In addition to the /s/-nasal clusters in (46), /sl/ and /sw/ are also affected:

(47) slide [ɬaid] (Stage 15)
sweet [w̥i:t/fi:t] (Stage 16)

When we look closer at A's data, we find that this phenomenon is not limited to /s/ clusters, but also occurs (though much less often) with target /f/-clusters:

(48) flapjack [læpdæk/ɬæpdæk] (Stage 15)
flower [læwə/ɬæwə] (Stage 15)

While space and scope limitations force me to leave this as an observational note, it does appear that the articulatory constraint proposed here is not directly involved. More likely, the feature [spread glottis] is important in these cases: the substitution begins soon after that feature is acquired at stage 13; and both fricatives are involved, not just /s/. It is possible that the asymmetry between the fricatives, with /s/-clusters more likely to show up as voiceless sonorants than /f/-clusters, is due in part to the articulatory constraint which targets /s/. However, it is difficult to tell if there is a true asymmetry here, or if it simply appears that way because there are fewer /f/-clusters in English to begin with.

## 6. Conclusion

The data presented in this thesis have posed some challenges to feature geometry theory if adopted to the exclusion of extralinguistic factors in phonological acquisition. The first challenge appeared in Chapter 3, when we saw that coronal stops were subject to velar but not labial harmony. Under Rice & Avery's (1995) geometry model, this asymmetry cannot be predicted. Instead, a model such as Brown's (1997) or Browman & Goldstein's (1986) proved necessary, in which [coronal] and [dorsal] are dependents of a single node opposing [labial] under the Place node. Such a model also allows for lingual articulations to form a natural class, proving useful in accounting for the problematic treatment of target /s/.

In Chapters 3 and 4, the substitution patterns of /s/ were looked at in detail. According to these patterns, /s/ behaved simultaneously like an underlying stop and an underlying continuant, depending on the featural environment. The decisive feature in the environment, it turned out, was [labial] as opposed to [coronal] and [dorsal]. Using Brown's model made it possible to capture this important conditioning factor since the two halves of the opposition ([labial] and [lingual], respectively) are directly represented in the tree. However, to fully capture the behavior of /s/ it was necessary to postulate an articulatory constraint against lingual continuants. Motivated by non-linguistic factors (the need for fine control of the tongue muscles), this constraint was shown to have linguistic consequences which gradually decreased as acquisition progressed. In the early stages, it effected the delinking of [continuant] from segments also specified [lingual], neutralizing the stop-fricative contrast for lingual segments only. Later, however – specifically, when fricatives began to receive a [spread glottis] specification – the constraint relaxed enough to allow [continuant] to exert a greater influence the production of target /s/. At this point, /s/ ceased to be realized as a plain stop; now the realization alternated freely between an aspirated and an affricated stop. Thus, the neutralized contrast essentially became narrower, involving only aspirates and fricatives.

Towards the last recorded stages, the effect of the constraint was weaker still, no longer effecting any neutralization at all. Instead, we found isolated instances of /s/ either preceded by an excrescent stop, or, in a cluster, lengthened in an insignificant way (length is certainly not contrastive on cluster-internal fricatives).

The proposals in this thesis have relied upon the assumption that physiological factors can influence the linguistic system. This assumption, in spite of its long-standing history, has yet to be formalized in its relationship with feature geometry theory. To what extent can factors such as neuro-motor control influence the development of phonological representations? In this thesis, I have attempted to show that such an influence does exist, echoing what Goad (1997) stresses: while phonological theory must be able to correctly predict and account for acquisition phenomena, it is essential that the theory be grounded in independent, functional motivations. It may well be that some of these motivations are first, and most clearly, apparent in child language.

**References**

Archibald, John (ed.). 1995. *Phonological Acquisition and Phonological Theory.* Hillsdale, NJ: Lawrence Erlbaum.

Avery, Peter and Keren Rice. 1989. Segment structure and coronal underspecification. *Phonology* 6(2):179-200.

Bernhardt, Barbara, John Gilbert and David Ingram (eds.). 1996. *Proceedings of the UBC International Conference on Phonological Acquisition.* Somerville, MA: Cascadilla Press.

Blevins, Juliette. 1995. The syllable in phonological theory. In Goldsmith (ed.) pp.206-244

Browman, Catherine and Louis Goldstein. 1986. Towards an articulatory phonology. *Phonology Yearbook 3*:219-252.

Browman, Catherine and Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology 6*(2):201-251.

Brown, Cynthia. 1997. Acquisition of segmental structure: consequences for speech perception and second language acquisition. PhD dissertation, McGill University.

Carr, Philip. 1993. *Phonology*. London, Eng: Macmillan Press.

Chin, Steven B. and Daniel A. Dinnsen. 1992. Consonant clusters in disordered speech: constraints and correspondence patterns. *Journal of Child Language 19*:259-285.

Clements, G. N. 1985. The geometry of phonological features. *Phonology Yearbook 2*:225-252.

Clements, G. N. and Elizabeth Hume. 1995. The internal organization of speech sounds. In Goldsmith (ed.) pp.245-306.

Clements, G. N. and S. J. Keyser. 1983. *CV Phonology: A Generative theory of the syllable*. Cambridge, MA: MIT Press.

Demuth, Katherine and E. Jane Fee. 1995. Minimal words in early phonological development. Ms. Brown University and Dalhousie University.

Demuth, Katherine. 1996. Alignment, stress and parsing in early phonological words. In Bernhardt, Gilbert and Ingram (eds.) pp.113-125.

Dinnsen, D. 1996. Context-sensitive underspecification and the acquisition of phonemic contrasts. *Journal of Child Language 23*:57-79.

Fee, E. Jane. 1995. Segments and syllables in early language acquisition. In Archibald (ed.) pp.43-61.

Fee, E. Jane. 1996. Syllable structure and minimal words. In Bernhardt, Gilbert and Ingram (eds.) pp.85-98.

Fikkert, P. 1994. *On the acquisition of prosodic structure*. Leiden: HIL.

Gnanadesikan, Amalia E. 1995. Markedness and faithfulness constraints in child phonology. Ms. University of Massachusetts, Amherst.

Goldsmith, John (ed). 1995. *The Handbook of Phonological Theory*. Cambridge, MA: Blackwell.

Goad, Heather. 1997. Consonant harmony in child language: an optimality theoretic account. In S. J. Hannahs and Martha Young-Scholten (eds.) pp.113-142.

Halle, Morris and William Idsardi. 1995. General properties of stress and metrical structure. In Goldsmith (ed.) pp.403-443.

Hannahs, S. J. and Martha Young-Scholten (eds.). 1997. *Focus on Phonological Acquisition.* Philadelphia PA: John Benjamins.

Hanson, R. 1999. Fusion and the acquisition of s-nasal clusters. Ms. University of Calgary. To appear, University of Maryland linguistics department Undergraduate Essay Prize in Linguistics website.

Hanson, R. 2000. VOT in English onsets: Phonetics and phonology interact. Ms. University of Calgary.

Hirose, Hajime. 1999. Investigating the physiology of laryngeal structures. In W. J. Hardcastle and J. Laver (eds.), *The Handbook of Phonetic Sciences.* Malden, MA: Blackwell. pp.116-136.

Ingram, David. 1996. Some observations on feature assignment. In Bernhardt, Gilbert and Ingram (eds.) pp.53-61.

Ingram, David. 1976. *Phonological Disability in Children.* London: Edward Arnold.

Kaye, Jonathan, Jean Lowenstamm and Jean-Roger Vergnaud. 1990. Constituent structure and government in phonology. *Phonology* 7(2):193-231.

Kiparsky, P. 1993. Blocking in non-derived environments. In S. Hargus, and E. Kaisse (eds.), *Phonetics and Phonology, Volume 4: Studies in lexical phonology.* San Diego, CA: Academic Press. pp.277-313.

LaCharité, Darlene and Carole Paradis. 1993. Introduction: The emergence of constraints in generative phonology and a comparison of three current constraint-based models. *The Canadian Journal of Linguistics Special Issue: Constraint-Based Theories in Multilinear Phonology.* pp. 127-153.

Ladefoged, Peter and Ian Maddieson. 1996. *The Sounds of the World's Languages.* Cambridge, MA: Blackwell.

Lévesque, Claire. 1992. Hachismo, or the behavior of /s/ in the Spanish of Seville. Honors thesis, University of Calgary.

Levin, J. 1988. A Place for lateral in the feature geometry. Ms. University of Texas, Austin.

Macken, Marlys. 1995. Phonological acquisition. In Goldsmith (ed.) pp.671-696.

Macken, Marlys. 1996. Prosodic constraints on features. In Bernhardt, Gilbert and Ingram (eds.) pp.159-172.

McCarthy, John. 1988. Feature geometry and dependency: a review. *Phonetica* 43:84-108

Paradis, C. and J.-F. Prunet (eds.). 1991. *Phonetics and Phonology 2: The Special status of coronals. Internal and external evidence.* New York: Academic Press.

Piggott, G. 1992. Variability in feature dependency: The case of nasality. *Natural Language and Linguistic Theory 10*:33-77.

Piggott, G. 1993. The geometry of sonorant features. Ms. McGill University.

Pollock, K. 1983. Individual preferences: case study of a phonologically delayed child. *Topics in Language Disorders 3*:10-23.

Rice, Keren and Peter Avery. 1995. Variability in a deterministic model of language acquisition: a theory of segmental elaboration. In Archibald (ed.) pp.23-42.

Rice, Keren and Peter Avery. 1991. On the relationship between laterality and coronality. In C. Paradis and J.-F. Prunet (eds.) pp.101-124.

Salidis, J. and J. Johnson. 1997. The production of minimal words: a longitudinal case strucy of phonological development. *Language Acquisition 6* (1):1-36.

Slobin, D. I. 1973. Cognitive prerequisites for the development of grammar. In C. Ferguson and D. I. Slobin (eds.), *Studies in Child Language Development.* New York: Holt, Reinhart and Winston. pp.175-208.

Smith, N. V. 1973. *The Acquisition of Phonology: A case study.* Cambridge: Cambridge University Press.

Steriade, D. 1987. Redundant values. In A. Bosch, B. Need, and E.Schiller (eds), *Parasession on autosegmental and metrical phonology.* Chicago: Chicago Linguistic Society. pp.339-362.

Velleman, Shelley. 1996. Metathesis highlights feature-by-position constraints. In Bernhardt, Gilbert and Ingram (eds.) pp.173-186.

Yip, Moira. 1990. Two cases of double-dependency in feature geometry. Ms. University of California at Irvine.

Rebecca Hanson
Department of Linguistics
University of Calgary
2500 University Drive
Calgary, Alberta
T1N 1N4 Canada
rjhanson@ucalgary.ca

## Knowledge and Performance: An Examination of the Role of Explicit Linguistic Knowledge in L2 Phonological Acquisition

Jennifer Mah
University of Calgary

### Abstract

A major challenge facing adults acquiring a second language is the mastery of the second language (L2) phonology. Previous work by Han (1992) shows that even advanced students who have been deemed by native speakers to be fluent in their second language may have failed to completely master the L2 phonological system, resulting in an audible non-native accent. The present paper explores a possibility suggested by Han (1992): that explicit knowledge of the L2 phonology is required to successfully establish the appropriate phonological representations and achieve native-like performance with respect to the L2 phonology. Japanese utterances produced by a native speaker of English were recorded after four months of classroom exposure and examined for accuracy in timing control and spectral accuracy of long and short vowels, and timing control of geminate and singleton consonants. The results were then presented to the subject, and a second sample of utterances was recorded after a further two months of classroom exposure. Although the subject showed evidence of having established separate phonological representations for the Japanese length distinctions (as indicated by *t* tests), her performance was still distinct from that of a native speaker. Furthermore, the subject was not able to make use of the explicit knowledge gained from the results of the first round of recording, as the subject's performance did not show any significant change in the second round of recording. These results suggest that the knowledge of Japanese length contrasts may have been useful to the learner in establishing the appropriate phonological representations, but not in gaining control over the finer articulatory details of these in six months.

### 1. Introduction

The acquisition of a second language is an intriguing process, differing importantly from first language acquisition in the final state achieved by the learner: while children arrive at full native competence levels in their first language, the degree of success achieved by second language learners is highly variable. Furthermore, while children acquiring their first language (L1) do so in a seemingly automatic and effortless fashion, L2 students are often seen to struggle, and the end result very rarely ends up being native-like, most remarkably so in the domain of phonology (Scovel 1969, reported in Archibald and Libben 1995), resulting in difficulty perceiving non-native contrasts and an audible non-native accent (Archibald 1996, in O'Grady and Dobrovolsky 1996).

The universal success in L1 acquisition, given that this complex process is being performed by the immature cognitive systems of young children, has led researchers to posit Universal Grammar (UG), a system of principles and parameters that constrains the possibilities considered by a child to be a part of their L1, thus facilitating the task of acquisition: principles are defined as those linguistic universals that account for the remarkable similarities that exist cross-linguistically, while parameters are those linguistic variables, often described as "switches" or a set of options (O'Grady 1996, in O'Grady and Dobrovolsky 1996), which allow for cross-linguistic variation. UG thus establishes the patterns that are permissible in human language, the child then approaches the task of L1 acquisition as a selector, setting parameters in accordance with the language of the environment. Despite the observed universal success of L1 acquisition, there are cases of individuals who have failed to fully acquire their L1. The most notable of these is the case of Genie (Curtiss 1977, reported in Berko-Gleason 1997) a girl who was deprived of exposure to the linguistic environment until the age of 11, at which time she was recovered by social services. Although she received intensive therapy, her linguistic skills never developed to native-like competence. Cases like Genie's have been touted as evidence for the Critical Period Hypothesis, which states that UG must no longer be accessible to the learner once a certain age has passed, usually set at around 13 years of age (Lenneberg 1967; Penfield 1965; both reported in Archibald and Libben 1995); however, much debate has arisen with respect to the ages associated with the Critical Period Hypothesis and the relation to neurological development (Long 1990).

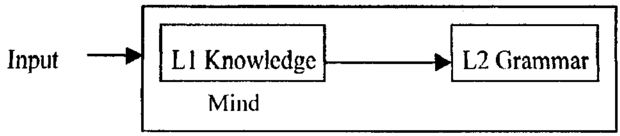
Further evidence in support of the Critical Period Hypothesis is found in L2 acquisition literature: L2 learners are found to rarely achieve native-like proficiency in the L2 (Newport 1990; Long 1990; Birdsong 1999). This body of research has prompted the proposal that UG is not accessible in L2 acquisition, formalized by Bley-Vroman (1989) as the Fundamental Difference Hypothesis: L1 and L2 acquisition differ in both the learner's definitions of a possible grammar, and in the procedure or set of procedures used to arrive at a grammar based on the available data (p. 51). Bley-Vroman (1989) argues that L2 acquisition proceeds by general problem solving skills which are not specific to the language faculty. On the other hand, White (1989) argues that UG does play a role in L2 acquisition, as some of the issues in L1 acquisition that prompted researchers to posit UG are also present in L2 acquisition. Both arguments do capture the fact that adult L2 learners' pronunciation is observably different from that of native speakers.

An interesting line of research has resurrected the notion that it is the L1, and not access to UG, that crucially affects competence in the L2. This transfer-based approach captures an important difference between L1 and L2 learners: only L2 learners come to the task of acquisition already knowing a language. Despite general agreement in the field that the L1 does play a role in L2 acquisition (Bley-Vroman 1989; White 1989), there is considerable debate as to how and to what extent the L1 affects L2 acquisition. Beginner L2 learners view the linguistic world as it corresponds to their knowledge of their first language; precisely how the previous knowledge affects the acquisition of new knowledge and the duration of this effect are still unknown.

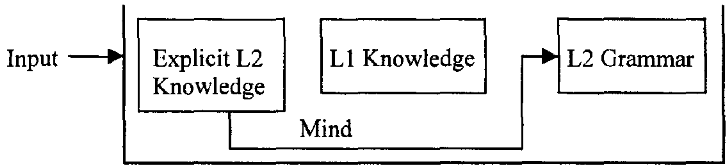
Furthermore, studies have shown that the filtering effects of the L1 are not insurmountable: there are cases of L2 learners who have been judged to have achieved a level of proficiency that is indistinguishable from the performance of native speakers (Bongaerts, Mennen and van der Slik 2000), and training in both perception and production of targeted segments has been shown to offer some improvement, albeit not to native-like proficiency (Matthews 1997). These studies, however, fail to address an important question: what is the factor that allows the learner to bypass knowledge imposed by the L1 and thus build the new structure required for L2 acquisition?

The present research aims to evaluate the possibility that it is explicit knowledge that is the key factor in determining the level of achievement possible in phonological acquisition for an adult L2 learner through examination of production data for evidence of acquisition of a novel phonological contrast. Here, the term explicit knowledge refers to the learner's conscious awareness of some linguistic property's presence and operation in the L2. The following diagrams may help to illustrate this point: (1) illustrates the input being modified by L1 knowledge in the mind of the learner before reaching the developing L2 grammar; (2) illustrates the possibility being examined here, that explicit linguistic knowledge intercepts the input before it is modified by the L1 knowledge, allowing the input to reach the developing L2 grammar as is.

(1)



(2)




The selection of L2 phonology as the domain to be examined is not accidental; a considerable body of research has been devoted to the study of the pronunciation of L2 learners. There are two reasons for this: firstly, as will be further discussed in Section 2 below, infants have been shown to demonstrate a sensitivity to their native language's phonological system that develops within the first year of life, both in terms of prosody (Mehler et al. 1998, reported in Jusczyk et al. 1995; Jusczyk, Cutler, and Redanz 1993) and segmental distinctions (Werker and Tees 1983, reported in Jusczyk 1997; Polka and Werker 1994); phonology may be the first linguistic domain to emerge through the course of L1 acquisition. Secondly, and perhaps more importantly given the context of this research, there is evidence that phonological abilities are the most sensitive to age effects. Success rates for acquisition of L2 phonology begin to decline as early as age 6 (Long 1990). These findings suggest that if L1 knowledge affects the perception and/or production of the L2 being acquired, its effects would be most observable in the L2 learner's phonology.

The paper is structured as follows: Section 2 will discuss the nature and implementation of the filter itself, through a survey of research dealing with infant perception, phonological development, and L2 learner performance; Section 3 will examine the contrasts being examined in the present research through a discussion of the relevant features of the L1 and L2 phonological systems; Section 4 will outline the methodology used to gather and analyze the data; Sections 5 and 6 will present the data collected and resulting analyses; Section 7 will sum up with a discussion of the results, their implications for both phonological theory and L2 acquisition theory, and directions for further research.

## 2. Linking L1 and L2 Acquisition

### 2.1. Early Phonological Development

Although children do not begin producing coherent, comprehensible utterances until they have reached the age of approximately two years, an extensive body of research has been devoted to the investigation of pre-verbal linguistic abilities (Burnham 1986; Friederici and Wessels 1993; Goodsitt, Morgan and Kuhl 1993; Hohne and Jusczyk 1994; Jusczyk 1993a, 1997; Jusczyk and Aslin 1995; Jusczyk, Cutler and Redanz 1993; Jusczyk, Hohne and Mandel 1995; Polka and Werker 1994). A variety of innovative investigative techniques have been developed that allow researchers to observe pre-verbal infants' linguistic behaviour by measuring their physical reactions to a change in the linguistic environment. Techniques such as the head-turn procedure, the preferential looking task, and the high-amplitude sucking procedure allow us to reliably assess the perceptual abilities of infants who have not yet begun to speak.

These studies of infants' perceptual abilities show a definite developmental path: at earlier stages of development, perceptual abilities can be termed as being language-general, in that they are able to discriminate a variety of contrasts, both native and non-native; conversely, at later stages of development (still in the first year of life) their discriminatory abilities are limited to those contrasts found in the L1 being acquired, which can be termed as language-specific abilities. Shortly after birth, infants have been shown to exhibit a high degree of sensitivity to the prosodic structure of the L1, to which they have been amply exposed in utero: Mehler et al. (1988, reported in Jusczyk, Hohne, and Mandel 1995) demonstrated that 4-day old infants were able to distinguish utterances in their mother's L1 from utterances in other languages. Furthermore, these same infants were unable to distinguish between utterances of two languages when neither was the mother's L1. Mehler et al.'s (1988) argument was that these discriminatory abilities were based on prosody; the wall of the uterus acts as a low-pass filter, allowing only the lower frequencies to reach the infant, thus the newborn has had ample experience with the prosodic structure of the L1 and is able to make the observed discriminations. More important to the present research, however, is the development of perceptual abilities with respect to segmental contrasts in the L1. Hohne and Jusczyk (1994) found that 2-month old infants displayed sensitivity to a variety of allophonic distinctions, and argue that it is unlikely that these very young infants are analyzing the segments as allophones, but rather are able to distinguish them since the distinctions may later become important phonological features used in setting up the L1's phoneme inventory and segmenting the speech stream. Jusczyk (1997) reports on a number of studies that provide evidence that infants are initially able to distinguish both native and non-native contrasts for a variety of segments, including novel contrasts. These infants are thus equipped with the perceptual abilities to acquire the phonological system of any of the world's languages.

Yet these language-general abilities do not remain, and, as was demonstrated in the case of prosodic structure, infants' perceptual abilities with respect to segmental structure follow a developmental path whereby the infant becomes attuned to the phonological contrasts and features of the L1 alone. Jusczyk (1997) reports on a number of investigations finding that a decrease in perceptual sensitivity to non-native consonant contrasts began setting in at 8 to 10 months of age, and was firmly in place by 10 to 12 months. In their examination of infants' perception of non-native vowel contrasts, Polka and Werker (1994) found that the decrease in perceptual abilities detected for non-native consonant contrasts set in earlier for non-native vowel contrasts: at 6 to 8 months of age, a decrease in sensitivity was observed, and this decrease was even more apparent among 10- to 12-month olds. The earlier onset of decrease in perceptual abilities for non-native vowels was argued to be due to the greater saliency of these segments in the speech stream. These studies demonstrate that within the first year of life, the infant's perceptual abilities undergo a massive developmental change, the result being that only those contrasts active in the language being acquired are recognized.

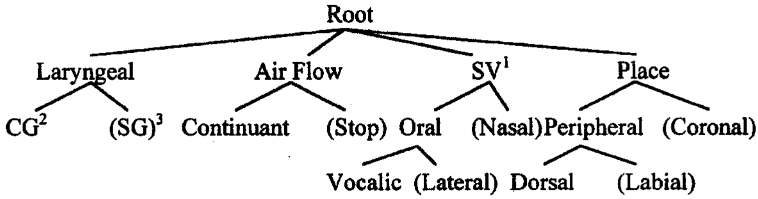
Researchers argue that this decline in perceptual abilities is beneficial to the task of L1 acquisition, as the system is appropriately constrained so as to reduce the processing load on the infant's immature cognitive system by restricting limited attention resources to the relevant contrasts in the environment (Jusczyk 1993). Yet it is precisely this developmental change in perceptual abilities that results in the difficulty L2 learners frequently have in perceiving and producing certain non-native contrasts. It appears that the filtering effects of L1 knowledge serve a beneficial purpose during L1 acquisition; yet this same knowledge serves a detrimental purpose through the course of L2 acquisition. This results in the L1 phonology's considerable influence on speech perception and the developing L2 system.

**2.2. Building Knowledge– A Model of Phonological Development**

Rice and Avery (1995) develop a theory of segmental elaboration based on feature geometry and present a model under which acquisition proceeds following pre-determined learning paths. Children begin the process of phonological acquisition with only phonological universals, the innate endowment from UG. The phonemic inventory of the L1 is built up through elaboration of structure: the child begins with an impoverished structure, and as phonemic contrasts are detected in the environment, the appropriate distinctive features are added to the grammar.

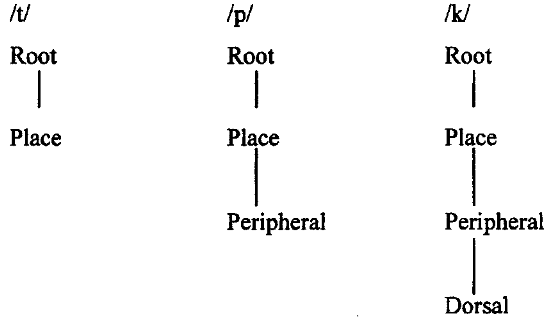
The notion of a learning path enters the picture when one considers how feature geometry organizes segmental structure. According to the feature geometry model, segments have internal structure, which is in turn organized into a hierarchical tree, with features grouped into nodes that reflect dependency and markedness relationships among the features. The basic tree given in (3) details the organization of the nodes and features that are used in describing segments using a model of feature geometry (Rice and Avery 1995: 31); unmarked features are enclosed in parentheses. These are not thought of as fully active in the phonology until they are motivated by an appropriate contrast; once a contrastive feature is added to the tree, then both features are specified.

(3)




Phonemic segments are underspecified; that is, redundant information (such as the unmarked features illustrated in (3)) is absent from the underlying phonological representations, as exemplified in (4) below. Given the contrast /t/ - /p/ - /k/ in (4), the step-by-step addition of distinctive features to the child's grammar allows development of the appropriate phonemic contrast.

(4)




(Rice and Avery 1995: 32)

Given the structures posited here, this model makes predictions about the order of acquisition of contrasts (i.e., children will first distinguish /t/ vs. /p/, followed by the three way contrast), which Rice and Avery (1995) claim are generally borne out in the existing data. Furthermore, given the dependency relations inherent in these structures, learning can only proceed along a predetermined path within any

---

[1] It should be noted that the abbreviation SV is used here to denote the Sonorant Voice node, which "organizes features normally associated with sonorant segments such as nasals, laterals, and r-sounds" (Rice and Avery 1995: 31). This does not refer to the feature of Voice, which, while not illustrated in the above tree, is associated with the Laryngeal node.

[2] The abbreviation CG is used to denote the Constricted Glottis node.

[3] The abbreviation SG is used to denote the Spread Glottis node.

particular organizing node. Differential expansion of phonemic inventories cross-linguistically is argued to be the result of elaboration within different nodes at of development; individual variation within a language is attributed to differential detection of (and importance attributed to) contrasts in the linguistic environment.

Thus we have seen evidence, from both observational studies and a theoretical standpoint, for filtering effects of L1 knowledge in the learner's perception of linguistic input (and potentially the learner's production output). Although they begin life with minimal phonological structure, enabling them to acquire any of the world's languages, the relevant segmental structure is elaborated so that the appropriate L1 contrasts are represented in the grammar. Children's perceptual abilities also become attuned to the L1 within the first year of life, severely reducing the ability to perceive the non-native contrasts that were perceptible at an earlier stage of development.

### 2.3. Filtering Effects of L1 Knowledge

Given the results from the studies reviewed above, an intriguing possibility presents itself to the researcher: there may be a causal relationship between phonological development and speech perception. We now turn to a study conducted by Brown (2000), in which she argues that the L1 filtering effects observed in both L1 and L2 acquisition are indeed a result of the segmental elaboration of feature geometry that takes place during L1 phonological acquisition. Brown (2000) claims that the acquisition of a phonological system not only determines the course of development for speech perception in L1 acquisition, but continues to constrain speech perception in L2 acquisition.

A survey of data from studies of categorical speech perception (reported in Brown 2000) suggests that the phonological system mediates between the acoustic signal and the linguistic system, allowing the speech stream to be broken down into meaningful units for processing and comprehension according to the phonemic category boundaries set out by the phonological system. This results in acoustically distinct stimuli being perceived as one segment, provided both tokens fall within the appropriate category boundaries. Thus the loss of perceptual abilities observed in studies of infant perception is not so much a loss, but a reorganization of those same perceptual abilities: the infant can still detect the variability of the acoustic signal, but has developed phonemic categories in which numerous tokens are deposited and interpreted as one distinct segment. That is to say, the system no longer attributes any linguistic importance to such variability in the speech stream.

This sort of perceptual reorganization has readily identifiable benefits, not only for first language acquisition but also for processing of language among adult native speakers: irrelevant "noise" occurring in the acoustic signal as a by-product of everyday language use, such as variable realizations, inter-speaker variability, and coarticulations, are disregarded, thus reducing the listener's processing load. Yet it quickly becomes apparent that this link between phonological development and speech perception acts as an impeding force that the listener must overcome in order to successfully acquire the L2 phonological system: it is fully possible that two phones used contrastively in the L2 may be interpreted as two separate instances of one phoneme in the L1, in which event the learner will be unable to perceive the contrast, such as Japanese learners' inability to differentiate between English /l/ and /r/, owing to the fact that Japanese has only one liquid phoneme (Matthews 1997). Effectively, Brown's (2000) proposal suggests that the phonological system of the learner's L1 will accurately predict which non-native contrasts will be perceived and thus successfully acquired, and which ones will not.

This same concept of L1 influence in L2 acquisition was the driving force behind earlier studies involving contrastive analysis, in which the phonemic inventories of the L1 and L2 were compared; the learner was expected to struggle with L2 contrasts involved phonemes that were absent from the L1 (Lado 1957, reported in Archibald and Libben 1995). This prediction, however, was not borne out: not only did L2 learners experience difficulty on certain novel contrasts, as was expected, they also experienced ease on other novel contrasts. This posed a rather hefty problem for the theory of contrastive analysis. Wardhaugh (1970, reported in Archibald and Libben 1995) attempted to compensate, by positing two versions of contrastive analysis: strong (or predictive), which attributes all areas of difficulty in the L2 to the structures of the L1 (this is the version originally set forth by Lado 1957), and weak (or diagnostic), which acknowledges that areas of L2 difficulty may stem from sources other than the L1. Brown (2000) claims that the failure of contrastive analysis lies in its identification of the components of the L1 grammar that cause interference: it is not the phonemes themselves, but rather the features contained in the learner's grammar that accurately predict perception of non-native contrasts. The featural level, rather than the segmental level, plays the crucial role in perception of L2 contrasts; therefore a feature-based approach is more appropriate in accounting for L2 phonological errors, rather than a segment-based approach. This is most clearly exemplified by Brown's (2000) research, which aimed to investigate the influence of Japanese, Korean, and Mandarin Chinese grammars on the acquisition of English contrasts not found in the L1.

Subjects participating in the study were native speakers of either Japanese, Korean, or Mandarin Chinese. The contrasts examined were /l/ - /r/, /b/ - /v/, /p/ - /f/, /f/ - /v/, and /s/ - /ʃ/ (another contrastive pair, /p/ - /t/, was included as a control item, as this contrast is active in all three L1s observed). These particular contrasts were chosen not only for their absence (by absence we are referring to an absence of the contrast, not of the phones themselves) in the native grammars of all subjects, but also for the variety of features upon which this group of contrasts relies. Table 1 below illustrates the various contrasts under investigation, and their status in the subjects' L1.

*Table 1*

| Contrast | Distinguishing Feature | Feature present in | | |
|---|---|---|---|---|
| | | Japanese | Korean | Mandarin Chinese |
| /l/ vs. /r/ | [coronal] (/r/) | no | no | yes |
| /b/ vs. /v/ | [continuant] (/v/) | yes | yes | yes |
| /p/ vs. /f/ | [continuant] (/f/) | yes | yes | yes |
| /f/ vs. /v/ | [voice] (/v/) | yes | yes | yes |
| /s/ vs. 7 / | [distributed] (7 /) | no | no | no |

(adapted from Brown 2000)

Thus, if the features contained in the L1 grammar are indeed responsible for the observed L1 influence, then the speakers should differ according to native language with respect to their performance on the noted English contrasts. The predictions of the model are as follows: all subjects are expected to perceive and acquire the /p/ - /f/, /b/ - /v/, and /f/ - /v/ contrasts, as the L1 grammars all contain the features relevant to these pairs; none are expected to be able to perceive the /s/ - 7 / pair, as none of the L1 grammars contain the relevant feature, thus this contrast is not expected to be acquired. The /l/ - /r/ pair, however, is expected to pattern differently: the Japanese and Korean speakers are not expected to be able to perceive this contrast, but the Mandarin Chinese speakers are, as their L1 grammars do contain the relevant feature. Contrastively, a segment-based approach would make the same predictions as the feature-based approach save the last one, in which case the segment-based approach predicts that all three groups will experience the same difficulty with the /l/ - /r/ pair, as none of the phonemic inventories contain both of these segments.

A 4IAX Discrimination Task was used to assess perception of the targeted contrasts: in each trial, subjects were presented with two non-word pairs involving the targeted segments, one of which would consist of a minimal pair (e.g., ra/ra vs. ra/la), they were then required to indicate which pair they believed consisted of two different non-words. Accurate performance in selecting the minimal pair as the instance of different non-words was taken to be indicative of ability to perceive the targeted phonological contrast. The results provided strong support for the proposed model. None of the groups were able to perceive the /s/ - 7 / pair. Only the Mandarin Chinese speaking group was able to perceive the /l/ - /r/ contrast, despite its absence in the L1. Performance on those contrasts which are, according to the model, acquirable (but are not already present) was also shown to be initially poor, due to inexperience with the contrast, but improved drastically with time so that the performance of the more advanced learners on this task was at native-like levels of proficiency: since the L1 grammar contained the feature required to build the appropriate phonological representations for those segments, the learners were able to eventually detect the contrast in the speech stream. It would seem that in the earliest stages of L2 acquisition, the sounds of the L2 are mapped directly onto the phonological structures of the L1 in any way that fits (no matter how poorly, as the case may be). As acquisition proceeds, mismatches that occur in the early direct mapping prompt a re-evaluation of the acoustic properties of the segment(s). If the appropriate features are present in the L1 grammar, then new phonological representations can be built to better accommodate the new segments, the result being L2 phonological acquisition.

**2.4. Additional Supporting Data – An Extension to Prosodic Phonology**

Although the model presented in Section 2.3 above was supported by evidence from segmental phonology, it is reasonable to expect that it would extend to the domain of prosodic phonology: both domains exhibit the developmental perceptual attunement described earlier in Section 2.1, although no data has yet been presented demonstrating that infants are initially able to distinguish a variety of non-native prosodic parameters, as exposure to L1 prosody begins before birth, making it extremely difficult to observe and assess the initial state: even newborn infants exhibit attunement to the L1 (Mehler et al. 1988). Still, an examination of L2 learners' behaviour with respect to prosodic phonology would serve as an appropriate test for Brown's (2000) model. Since both segmental and prosodic phonology seem to exhibit the same developmental course in L1 acquisition, we would want to posit one mechanism to account for both, and for the resulting structures in the grammar and their operation in L2 acquisition. An earlier study by Han (1992) provides data with respect to such an extension of Brown's (2000) model.

Han (1992) examined native speakers of English with respect to their performance on Japanese geminate stops. The study looked at the performance of 4 native speakers of American English who had been judged (by a native Japanese speaker) to be fluent in Japanese on a production task designed to elicit the Japanese single – geminate stop contrast. Measurements were taken of the subjects' timing control of the stop closure, and these were compared to the performance of native speakers of Japanese on the same task.

According to Brown (2000), the English speakers, no matter how advanced their competence in Japanese, should not be able to produce the single – geminate stop contrast, since the feature for consonant length is not active in their L1 grammars (it should be noted here that some researchers have argued that English does make use of geminate consonants in connected speech, as in *bla<u>ck</u>*

*cat*; these, however, are analyzed as phonetic, and not phonemic). Han's (1992) data do support this prediction, finding that the English speakers either missed the contrast altogether, or consistently produced one that was not native-like: while native Japanese speakers have been shown to produce a geminate – single stop mean timing ratio of approximately 3:1 (Han 1992; Homma 1981; Han 1962), the English speakers produced a mean timing ratio of approximately 2:1 (Han 1992). Table 2 below details the mean stop closure timing control ratios produced by native Japanese speakers; Table 3 describes the mean stop closure timing control ratios of the American subjects, broken down by subject (Han 1992). The graph in (5) below illustrates the difference between the mean ratios produced by native Japanese speakers and the mean ratios produced by Han's non-native speakers.

*Table 2: Native Japanese Speakers*

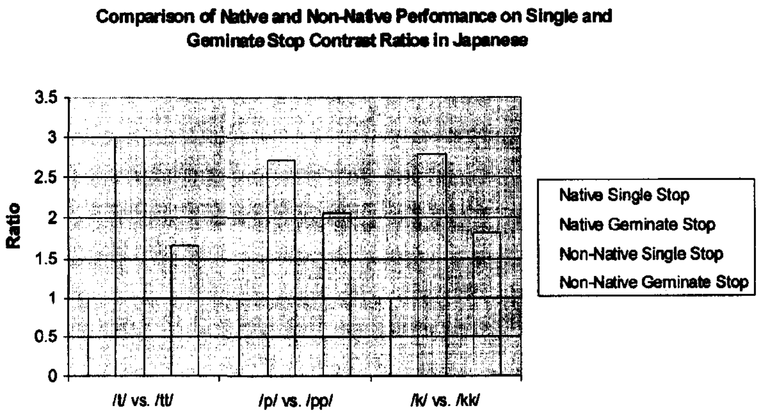
| | /tt/ vs. /t/ | /pp/ vs. /p/ | /kk/ vs. /k/ |
|---|---|---|---|
| Mean Ratio | 3.00 | 2.71 | 2.80 |

*Table 3: American Speakers of Japanese as a Second Language*

| Subject | [tt] vs. [t] 1 | [tt] vs. [t] 2 | [tt] vs. [t] 3 | [tt] vs. [t] 4 | [tt] vs. [t] 5 | [pp] vs. [p] 1 | [pp] vs. [p] 2 | [kk] vs. [k]1 | [kk] vs. [k] 2 | [kk] vs. [k] 3 |
|---|---|---|---|---|---|---|---|---|---|---|
| A | 1.37 | 1.55 | 2.10 | 1.48 | 1.26 | 2.03 | 1.73 | 1.45 | 2.27 | 1.71 |
| B | 2.57 | 3.35 | 2.76 | 2.97 | 2.27 | 3.56 | 2.75 | 2.55 | 4.02 | 1.94 |
| C | 1.21 | 1.85 | 1.01 | 1.01 | 1.50 | 2.40 | 1.89 | 1.75 | 1.65 | 1.08 |
| D | 1.00 | 1.18 | 0.90 | 1.06 | 0.93 | 1.10 | .098 | 0.99 | 1.22 | 1.14 |
| Mean | 1.54 | 1.98 | 1.69 | 1.63 | 1.49 | 2.27 | 1.84 | 1.68 | 2.29 | 1.47 |
| | 1.67 | | | | | 2.06 | | 1.81 | | |

(adapted from Han 1992: 118)

(5)




## 2.5. Remarks

We have seen how recent developments in language acquisition research have led to the proposal of a model that links L1 phonological acquisition to speech perception, and the related difficulties encountered in the acquisition of L2 phonology. In surveying the previous work, however, two interesting questions are raised. First, Brown's (2000) model makes predictions about the English speaker's control of length for both consonants and vowels in Japanese: we would expect the early learner to perform well with respect to vowel length, since English vowels are specified as being either mono- or bi-moraic (more on this in Section 3 below), all the while performing poorly with respect to consonant length, even at an advanced stage of acquisition, since English consonants are not specified for length. Second, Brown's (2000) model offers a rather bleak view of the final state of L2 acquisition: it can never be completely native-like if a relevant feature is missing from the L1 grammar. Yet Bongaerts, Mennen, and van der Slik (2000) report on a number of non-native speakers of Dutch from a variety of L1 backgrounds (German, English, French, Spanish, Armenian, Berber, Czech, Greek, Italian, Swedish, and Turkish) who have achieved a level of proficiency with respect to pronunciation that has been deemed, by native speakers of Dutch, to be indistinguishable from that of a native speaker. Clearly it is possible to overcome the filtering influence of the L1; the question that remains is how. Han (1992) suggests that explicit instruction may be the key to improving the learner's acquisition of Japanese length contrasts. The present research aimed to address both of these questions by evaluating a native speaker of Canadian

English's performance on Japanese consonant and vowel length contrasts; the subject was provided with various types of explicit knowledge throughout the investigation. Before moving on to our discussion of the experiment and its results, a discussion of the relevant phonological structures in both English (the L1) and Japanese (the L2) is in order.

## 3. Comparing the L1 and L2 Phonology

The present study sought to examine performance on Japanese length contrasts by a native speaker of Canadian English. The following discussion of English and Japanese phonology will elucidate the reasons behind the choice of this particular set of contrasts.

Let us start with the phonemic inventories of the L1 and L2. Tables 4 and 5 below detail the consonant inventories for English and Japanese, respectively; Tables 6 and 7 below detail the vowel inventories for English and Japanese, respectively.

*Table 4: Consonant Phonemes of English*

| | | Bi-labial | Inter-dental | Al-veolar | Palato-Alveolar | Palatal | Velar | Glottal |
|---|---|---|---|---|---|---|---|---|
| Stops | Voiceless | p | | t | | | k | |
| | Voiced | b | | d | | | g | |
| Fricatives | Voiceless | f | ·· | s | ·· | | | h |
| | Voiced | v | ·· | z | j | | | |
| Affricates | Voiceless | | | | ṫ | | | |
| | Voiced | | | | d̈ | | | |
| Nasals | | m | | n | | | ·· | |
| Liquids | Lateral | | | l | | | | |
| | Rhotic | | | r | | | | |
| Glides | | (w)[4] | | | | j | w | |

(adapted from O'Grady and Dobrovolsky 1996: 32)

[4] The placement of /w/ in parentheses here is due to its classification in English as a labio-velar, having properties of both labials and velars.

*Table 5: Consonant Phonemes of Japanese*

| | | Bilabial | Al-veolar | Alveolo-palatal | Palatal | Velar | Glottal |
|---|---|---|---|---|---|---|---|
| Stops | Voiceless | p pp | t tt | | | k kk | |
| | Voiced | b | d | | | g | |
| Fricatives | Voiceless | | s ss | ·· ·· | | | h |
| | Voiced | | z | | | | |
| Affricates | Voiceless | | | ṫ | | | |
| | Voiced | | | d̈ | | | |
| Nasals | | m mm | n nn | | | | |
| Liquids | | | ·· [5] | | | | |
| Glides | | | | | j | w | |

(adapted from Han 1962)

*Table 6: Vowel Phonemes of English*

| | Front | Mid | Back |
|---|---|---|---|
| High | ij ·· | | uw ·· |
| Mid | ej ·· | ·· | ow ·· |
| Low | ·· | | ·· |
| Diphthongs | aj, oj | | aw |

(adapted from O'Grady and Dobrovolsky 1996: 35)

*Table 7: Vowel Phonemes of Japanese*[6]

| | Front | Mid | Back |
|---|---|---|---|
| High | i i: | | u u: |
| Mid | e e: | | o o: |
| Low | | a a: | |

(adapted from Han 1962)

From these tables, we can see clearly that length is a contrastive feature in Japanese for both consonants and vowels, but only for vowels in English. Furthermore, English has a larger vowel inventory with respect to spectral quality of vowels, as well as three diphthongs which are absent in the Japanese vowel

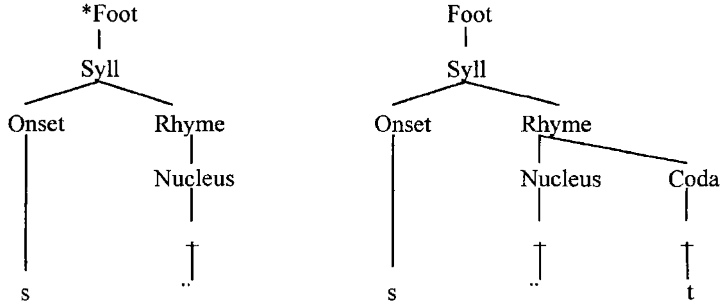
[5] The use of [ ] to represent the alveolar lateral flap is based on Rogers (1991), p. 228.

[6] Homma (1973) transcribes /u/ and /u:/ as 7 / and 7 :/ respectively, with some instances being recorded phonetically as [ ] and [ :]. It should be noted that the subject also tends to perceive these vowels as being unrounded; additionally, /o/ and /o:/ are perceived as 7 / and 7 :/. Thus, 7 /, 7 :/, 7 /, and 7 :/ are the transcriptions that will be used throughout the remainder of the paper.

inventory. Thus, it is of immediate interest to examine the learner's performance not only in terms of length contrasts (for both consonants and vowels), but also in terms of vowel quality.

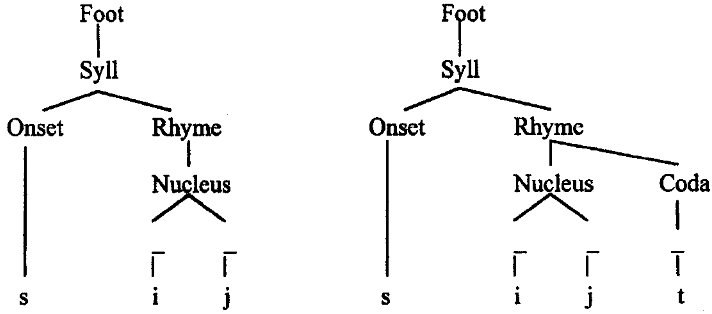
Japanese is a language in which moras are the units of prosodic significance (Han 1994, Port et al. 1987). Moras measure out syllable weight: a light syllable consists of one mora, a heavy syllable consists of two (or more) moras (Spencer 1996). Thus, Japanese short vowels are analyzed as being mono-moraic, while long vowels are bi-moraic. The analysis of English vowels is similar: in English, two classes of vowels can be identified: tense vowels (/ij, ej, uw, ow,¨ /) and lax vowels (7 ,¨ ,¨ ,¨ ,¨ ,¨ /), of which tense vowels are all inherently longer than the corresponding lax vowels. This has been analyzed as a phonemic distinction, resulting in an analysis of English lax vowels as being mono-moraic, while English tense vowels are bi-moraic. Evidence for this analysis comes from examination of English mono-syllabic feet, which must be bi-moraic, as illustrated by (6) below: a mono-moraic lax vowel alone may not form the rhyme of a mono-syllabic foot (*si); however, once a consonant fills the coda position (and is consequently assigned one mora), legal foot structure is obtained (*sit*). Conversely, a bi-moraic tense vowel may form the rhyme of a mono-syllabic foot, both with (*seat*) and without (*see*) a Coda consonant, as illustrated in (7).[7] (Spencer 1996)

(6)




[7] This analysis of vowel length in English is not without controversy; Dobrovolsky (in discussion) notes that English vowel length has been argued to be a phonetic feature resulting from the Advanced Tongue Root distinctive feature in the grammar: vowels that are specified as [-ATR] cannot be long.

(7)




Han (1992) argues that it is not the actual durational values obtained through measurement that will provide the accurate picture of the durational contrast, as these are easily confounded by differences in rate of speech and speaker-to-speaker variation; rather, it is the ratio obtained from the mean measurements that will provide a meaningful illustration of the nature of the contrast. The ratio of geminate to single stop consonants, illustrated in Section 2.4 above, is approximately 3:1 (Han 1992, Homma 1981, Han 1962), while the ratio of long to short vowels is between 2:1 and 3:1; Han (1962) claims that standalone vowels are produced using a ratio of 2:1, vowels preceded by a voiced consonant are produced using a ratio of 2.5:1, and vowels preceded by a voiceless consonant are produced using a ratio of 3:1. (It should be noted here that due to vocabulary limitations, the consonantal environment of the vowel was not taken into account in this study.)

Given the phonological descriptions presented above, Brown's (2000) model would make the following predictions about L2 performance:

1. A native speaker of English will perform poorly on Japanese consonant length contrasts, even at advanced stages of acquisition.
2. A native speaker of English will perform well on Japanese vowel length contrasts in more advanced stages of acquisition, but may perform poorly in early stages.
3. Substitution of English vowels may occur in early stages as the learner's developing interlanguage grammar attempts to cope with the new vowel system, but this should disappear at later stages of development.

Remembering Han's (1992) proposal that explicit instruction may assist the learner in acquiring Japanese consonant length contrasts, we can make the following additional prediction:

4. A native speaker of English who has received explicit instruction on Japanese consonant and vowel length contrasts will perform well on these, even at early stages of acquisition.

Indeed, any data supporting Han's (1992) proposal would suggest that linguistic knowledge gained through explicit instruction is available to the developing L2 grammar, a topic still under considerable debate.

## 4. Research Methodology

### 4.1. Subject

Only one subject (the researcher) provided data for this investigation: a 22-year old female student at the University of Calgary whose native language is Canadian English. The speaker has been a resident of Calgary, Alberta, for the past ten years, prior to that she resided in Edmonton, Alberta. In addition to English, this speaker is also fluent in French and speaks some Spanish: she began acquiring French at the age of five through a French immersion education program, and she began acquiring Spanish at the age of 18 when she began her studies at the university. Although she is not monolingual, her competence with these other languages is not expected to affect her acquisition of Japanese due to the large typological differences that have been observed; particularly, neither French nor Spanish makes use of length contrasts for either consonants or vowels.

The subject began acquiring Japanese in a university classroom setting at the age of 22. Classes were held four times a week, for one hour each class, with an additional hour per week spent on drill exercises that emphasized grammatical use of the linguistic structures of the Japanese language. Very little attention was given to pronunciation; the instructor made minimal comments about long vowels (referring to them as "stretched"), and even fewer comments about geminate consonants (describing the production of these as "swallowing the first part of the sound"). Prior to beginning classes, the subject's linguistic knowledge of Japanese was extremely limited: due to exposure to examples from various Linguistics classes, she knew that Japanese had a five vowel system (in terms of vowel quality), long and short vowel contrasts, and single and geminate consonant contrasts; however, any other knowledge of the language or its operation was nonexistent.

### 4.2. Data Collection

Data was collected in two sessions, two months apart: once after approximately four months' classroom exposure, and a second time after approximately six months' classroom exposure. Fifteen Japanese sentences were designed to elicit tokens of the targeted contrasts, albeit not in minimal pairs due to vocabulary restrictions. These sentences were each written on an index card using the *hiragana* Japanese script, which the subject had been required to master after one month of classes as part of the course requirements, so that attention and processing abilities would be devoted to decoding meaning from the *hiragana* script, rather than focussing on pronunciation. The cards were randomized, then presented to the subject for reading. To avoid unnatural pauses (and any lengthening that may have resulted) in the reading, only words that were present in the subject's working vocabulary were used; hence the absence of minimal pairs in the data. Another unfortunate result of this vocabulary familiarity condition is that a small number of tokens containing geminate consonants and/or long vowels were deemed appropriate for elicitation. It should be noted here that an additional sentence was presented in the second round of recording to elicit tokens of a geminate consonant (/ss/) that were not deemed appropriate to the first round of recording due to the aforementioned vocabulary restraint. The test sentences are listed in Appendix A.

The subject read each sentence three times, each of which was recorded using a Sony TCD – D100 Digital Audio Tape (DAT) recorder and a Sony ECM – MS908C electret condenser microphone. The recorded data were then re-digitized at a sampling rate of 22.2 kHz using the SoundScope 8 One Channel Analyzer in the Phonetics Laboratory at the University of Calgary. Wide-band spectrograms were made of the recorded data using this same device.

### 4.3. Measurement Procedures

The following measurements were all obtained through examination of the wide-band spectrograms obtained from the recorded data. The segments under scrutiny were isolated visually using the spectrogram, and then played back to ensure the absence of neighbouring sounds.

#### 4.3.1. Consonant Closure Duration

Stop closure duration was measured by the absence of noise on the spectrogram; fricative closure duration was measured from onset to endpoint of the characteristic "noise" on the spectrogram; nasal consonant duration was measured from onset to endpoint of the characteristic "nasal murmur" on the spectrogram. Due to the difficulty in determining the closure duration of stops in initial position, only intervocalic stops were included in the data analysis. Similarly, since fricatives in final position have a tendency to decrease in amplitude slowly, making it difficult to determine the actual endpoint for these sounds, these were also excluded from analysis.

#### 4.3.2. Vowel Duration

Vowel duration was measured from the onset of glottal vibration, indicated by regular vertical striations on the spectrogram, to the following closure. In the case of vowels in final position, the endpoint of the vowel was taken to be where glottal vibration ceased.

#### 4.3.3. Vowel Quality

Vowel quality was measured in terms of the values of the first and second formants (F1 and F2, respectively) of the vowel, which contain the most information with respect to that vowel's identity in acoustic space (Borden, Harris, and Raphael 1994). F1 and F2 values are easily obtained from a wide-band spectrogram.

#### 4.3.4. Remarks

In addition to being examined for significant differences in duration between single and geminate consonants and short and long vowels, all durational measurements were compiled into mean ratios, and compared against those obtained from studies of native speakers' performance on production of the targeted contrasts, as well as Han's (1992) results involving non-native speakers. Vowel quality measurements were compared against averaged measurements of a female native speaker's productions (Homma 1973) as well as measurements of the subject's own production of English vowels. Additionally, the results from Time I were compared against those from Time II in order to identify any changes in the subject's performance.

## 5. Results – Time I

### 5.1. Consonant Closure Duration

The graph in (8) below illustrates the subject's mean timing control of single and geminate consonants; Table 8 details the timing control ratios produced, as illustrated in (8).

(8)

Mean Japanese Consonant Duration as Produced by a Non-Native Speaker

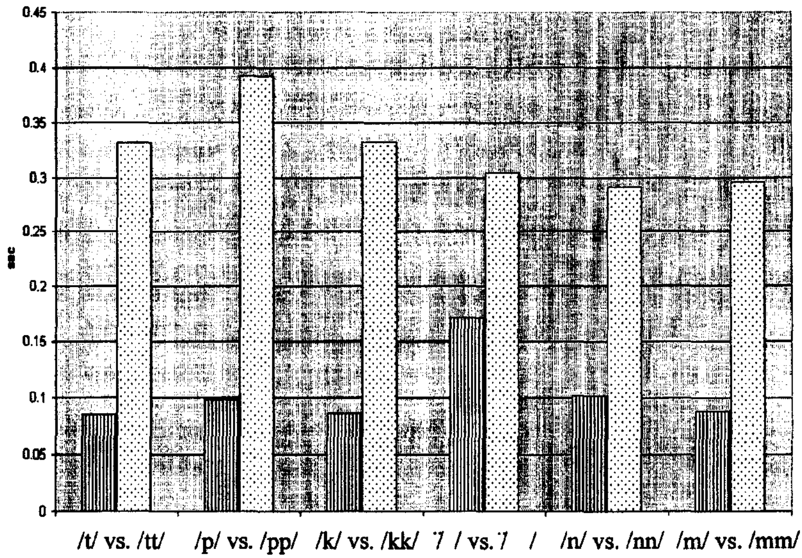


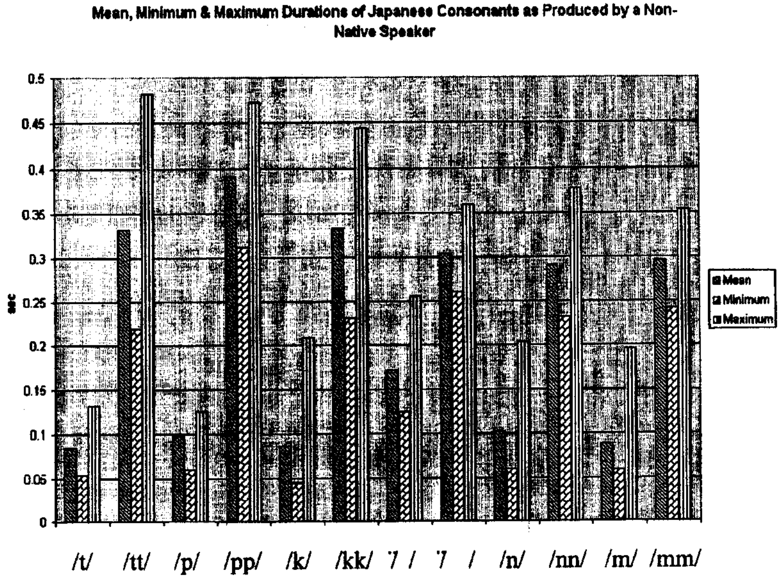

*Table 8: Mean Consonant Closure Duration Ratios*

| | Single | Geminate | Ratio |
|---|---|---|---|
| /t/ vs. /tt/ | 0.085 | 0.332 | 3.91 |
| /p/ vs. /pp/ | 0.098 | 0.392 | 4.00 |
| /k/ vs. /kk/ | 0.086 | 0.333 | 3.87 |
| '7 / vs.'7 / | 0.171 | 0.291 | 1.78 |
| /n/ vs. /nn/ | 0.102 | 0.291 | 2.85 |
| /m/ vs. /mm/ | 0.088 | 0.296 | 3.36 |

Overall, the subject produced a mean consonant closure duration ratio of 3.26:1.0. This appears promising for our hypothesis, given the mean overall closure duration of 2.8:1.0 (Han 1992) to 3:1 (Han 1962, Homma 1981) reported for native speakers. For all contrasts investigated, a two-tailed *t* test revealed that geminate consonants were significantly longer than their corresponding single consonants [/t/ vs. /tt/: $t = -12.820$, $p < 0.001$; /p/ vs. /pp/: $t = -10.649$, $p < 0.001$; /k/ vs. /kk/: $t = -7.292$, $p = 0.001$; '7 / vs.'7 /: $t = -8.044$, $p < 0.001$; /n/ vs. /nn/: $t = -7.640$, $p < 0.001$; /m/ vs. /mm/: $t = -12.657$, $p < 0.001$]. These results conflict with the predictions made by Brown's model of L1 interference, based on which we would expect poor performance on these contrasts. The subject's actual performance suggests that she has been able to acquire new phonological structure with respect to Japanese single and geminate consonants and overcome the L1 grammar's inadequacies in dealing with the novel contrast. These results also support our hypothesis that explicit linguistic knowledge is required to overcome the L1 influence: the subject was aware of consonant length contrasts in Japanese and was also able to consistently produce a significant consonant length contrast; at this point, however, it should not be expected that the subject's performance be fully native-like since she was not aware of the specific timing ratios involved.

The contrast she is producing, however, is not native-like in other important but troubling ways. Performance varies among the segment classes: among the obstruents, the ratios obtained for performance on stop contrasts are greater than those of native speakers, while the ratio obtained for the fricative contrast (/ / vs / /) is smaller than that produced by native speakers. Furthermore, closer inspection of the raw data reveals some variation in the tokens produced for each contrast: for some contrasts, the longest token of a single consonant was just as long or longer than the shortest token of the corresponding geminate consonant. The graph in (9) serves to illustrate this point. This variation both across and within categories suggests that although the subject reliably and consistently distinguishes between single and geminate consonants in her productions, she does not control the timing of these in the same way that a native speaker of Japanese would. Again, this is not surprising at this point, given that she was unaware of how her performance differed from that of a native Japanese speaker when the data were collected. If Han's (1992) proposal that explicit linguistic knowledge will enable the L2 learner to achieve native-like proficiency is correct, then we would expect to see this situation remedied at the second recording.

(9)




## 5.2. Vowel Duration

The graph in (10) illustrates the subject's mean timing control of vowel duration; Table 9 details the timing control ratios produced illustrated in (10).

(10)

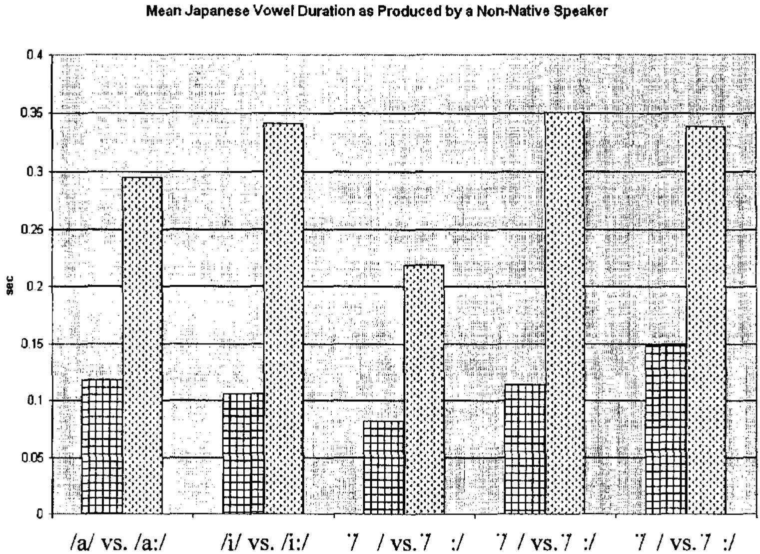


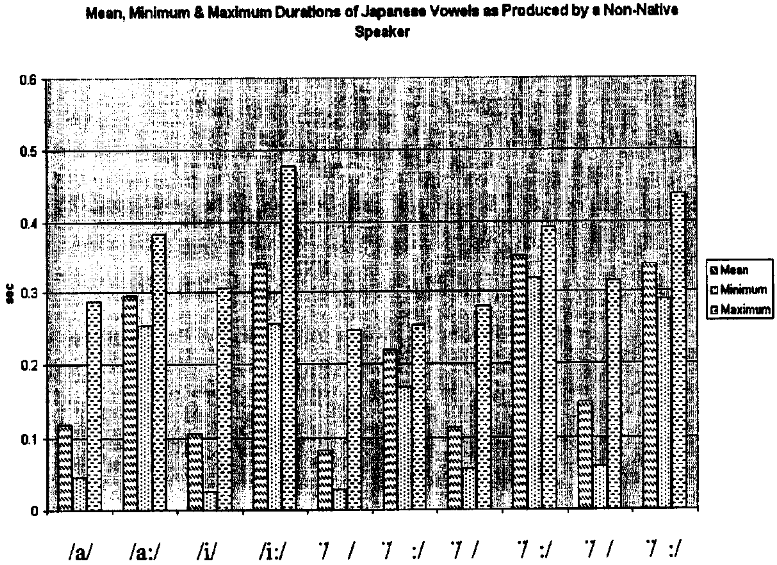

*Table 9: Mean Vowel Duration Ratios*

| | Short | Long | Ratio |
|---|---|---|---|
| /a/ vs. /a:/ | 0.118 | 0.295 | 2.50 |
| /i/ vs. /i:/ | 0.106 | 0.341 | 3.22 |
| 7 / vs.7 :/ | 0.082 | 0.219 | 2.67 |
| 7 / vs.7 :/ | 0.114 | 0.351 | 3.08 |
| 7 / vs.7 :/ | 0.148 | 0.339 | 2.29 |

Overall, the subject is producing a mean vowel duration contrast ratio of 2.65:1.0, which falls well within the reported native speaker range of 2:1 to 3:1 (Han 1962). For all contrasts investigated, a two-tailed *t* test revealed that long vowels were produced significantly longer than short vowels [/a/ vs. /a:/: $t = -9.716$, $p < 0.001$; /i/ vs. /i:/: $t = -5.818$, $p = 0.002$; 7 / vs. 7 :/: $t = -9.276$, $p < 0.001$; 7 / vs. 7 :/: $t = -12.668$, $p < 0.001$; 7 / vs. 7 :/: $t = -8.523$, $p < 0.001$]. This is consistent with predictions made by Brown's model, under which the subject was expected to have little difficulty in implementing the Japanese vowel length contrast since vowels in English are specified for length through the mono-moraic – bi-moraic distinction. These results are also consistent with our explicit linguistic knowledge hypothesis, as the subject was aware of the vowel length contrast in Japanese.

Again, while an examination of the mean ratios appears promising, an examination of the raw data reveals variation in the tokens produced, to a greater extent than that observed among the consonants: vowel duration is highly variable, in that for all vowels except'7 /, the longest token of the short vowel is as long or longer than the shortest token of the corresponding long vowel. The graph in (11) serves to illustrate this point.

(11)

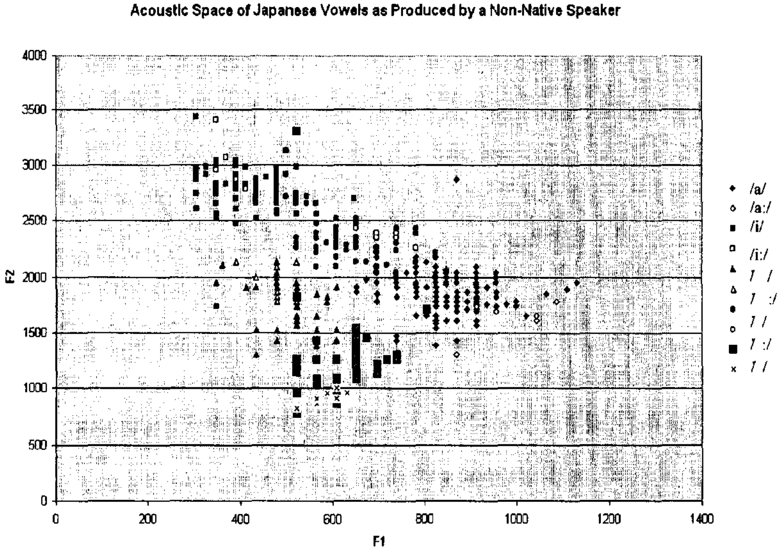



Although examination of the mean ratios yielded native-like results, this variability suggests that the subject is experiencing difficulty in controlling vowel length. Once again, this suggests a lack of native-like timing control of the L2 length contrast, a situation that our explicit linguistic knowledge hypothesis predicts would be remedied at the second recording.

## 5.3. Vowel Quality

The scatter plot presented in (12) details the subject's vowel productions in acoustic space.

(12)




The mean F1 and F2 frequencies are given in Table 10 below, as well as the mean F1 and F2 frequencies reported for a female native speaker of Japanese (Homma 1973).

*Table 10: Mean F1 and F2 of Japanese Vowels*

| | F1 - Native | F1 – Non-Native | F2 - Native | F2 – Non-Native |
|---|---|---|---|---|
| /a/ | 1046 | 855 | 2075 | 1870 |
| /a:/ | 1046 | 985 | 2075 | 1630 |
| /i/ | 354 | 413 | 2886 | 2783 |
| /i:/ | 354 | 376 | 2886 | 2991 |
| '7 / | 367 | 507 | 2060 | 1772 |
| '7 :/ | 367 | 473 | 2060 | 1941 |
| '7 / | 655 | 653 | 2209 | 2309 |
| '7 :/ | 655 | 717 | 2209 | 2361 |
| '7 / | 659 | 625 | 977 | 1208 |
| '7 :/ | 659 | 579 | 977 | 912 |

As is made clear by Table 10, the subject's Japanese vowels differed from those produced by a native speaker. Another important difference between native speaker performance and the subject's performance was noted: the subject produced significantly different mean vowel qualities for short vowels and their corresponding long vowels. Table 11 serves to summarize the results of statistic analysis. A two-tailed *t* test revealed that with the exception of the '7 / vs. '7 :/ contrast, all long vowels were pronounced with a significantly different quality than their corresponding short vowels[8]. This is clearly L1 interference, as English vowels are differentiated using both spectral and temporal cues, and these results serve as counterevidence to our explicit linguistic knowledge hypothesis, as it does not predict this spectral differentiation between long and short vowels. These results are in keeping, however, with Brown's (2000) model, which predicts that the subject will perform well with respect to the vowel contrasts at advanced stages of acquisition, but may exhibit difficulty at earlier stages.

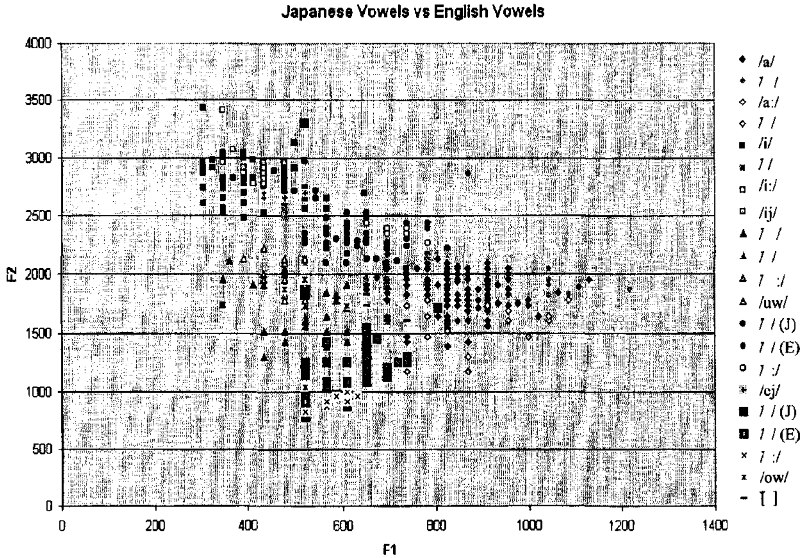
[8] Significance for vowel quality was determined on the following basis: if either or both of the F1 and F2 values reached significance, then the long vowel was deemed to be of a significantly different quality than its short counterpart. Since both the F1 and F2 are used in identifying the vowel in acoustic space, a significant change to either or both would result in the production of a different vowel.

*Table 11: Results of Two-Tailed t Test for Japanese Vowel Quality – Time I*

| Vowel Contrast | | t value | p value |
|---|---|---|---|
| /a/ vs. /a:/ | F1 | -3.825 | p < 0.001 |
| | F2 | 3.464 | 0.001 |
| /i/ vs. /i:/ | F1 | 2.860 | 0.016 |
| | F2 | -2.663 | 0.009 |
| 7 / vs. 7 :/ | F1 | 1.569 | 0.123 |
| | F2 | -2.285 | 0.027 |
| 7 / vs. 7 :/ | F1 | -1.795 | 0.077 |
| | F2 | -0.845 | 0.401 |
| 7 / vs. 7 :/ | F1 | 3.256 | 0.004 |
| | F2 | 2.504 | 0.015 |

The vowel quality data gathered were also compared against the F1 and F2 of Canadian English vowels to check for substitution of these in the subject's Japanese productions: Japanese short vowels were compared against English lax vowels; Japanese long vowels were compared against English tense vowels; this is illustrated in (13) below.

(13)



Results of this analysis are mixed, and do not wholly support or reject either of our hypotheses. While most Japanese vowels differed significantly from their English counterparts, the contrasts 7 / vs. 7 / and 7 :/ vs. /uw/ did not, suggesting that the subject is substituting rounded back vowels from English for the Japanese unrounded targets. This finding is consistent with Brown's model, which predicts that 7 / and 7 :/ should not be acquirable to our subject, as English high back vowels are not specified for roundedness (and are thus rounded by default); however, it also poses a serious problem for our explicit linguistic knowledge hypothesis, which claims that since the subject is aware of unrounded back vowels in Japanese, she should be able to produce them. Table 12 details the results of statistical analysis.

*Table 12: Results of Two-Tailed t Test for Japanese and English Vowels – Time I*

| Vowel Contrast | | t value | p value |
|---|---|---|---|
| /a/ vs. 7 / | F1 | -6.647 | $p < 0.001$ |
| | F2 | -0.509 | 0.611 |
| /a:/ vs. 7 / | F1 | 3.245 | 0.006 |
| | F2 | 1.111 | 0.287 |
| /i/ vs. 7 / | F1 | -5.309 | $p < 0.001$ |
| | F2 | 7.905 | $p < 0.001$ |
| /i:/ vs. /ij/ | F1 | -5.437 | $p < 0.001$ |
| | F2 | 0.997 | 0.362 |
| 7 / vs. 7 / | F1 | -2.114 | 0.083 |
| | F2 | 1.032 | 0.307 |
| 7 :/ vs. /uw/ | F1 | 0.923 | 0.373 |
| | F2 | -1.209 | 0.248 |
| 7 / vs. 7 / | F1 | -2.386 | 0.020 |
| | F2 | 4.161 | $p < 0.001$ |
| 7 :/ vs. /ej/ | F1 | 10.716 | $p < 0.001$ |
| | F2 | -11.939 | $p < 0.001$ |
| 7 / vs. 7 / | F1 | 2.673 | 0.010 |
| | F2 | 0.890 | 0.377 |
| 7 :/ vs. /ow/ | F1 | 2.603 | 0.022 |
| | F2 | -2.908 | 0.033 |

### 5.4. Remarks

Given that at the time of recording, the subject was aware only of the existence of geminate consonants and long vowels in Japanese, and not of her performance with respect to that of native speakers, the results of the analysis of duration are not surprising under our hypothesis that it is explicit linguistic knowledge that allows the learner to bypass L1 knowledge in building new structure for the L2. The subject appropriately produced geminate consonants and long vowels where required; however, some troubling variation in duration was found, particularly among the vowels.

The results of the analysis of vowel quality are intriguing, as the vowels produced neither wholly resembled those produced by native speakers of Japanese, nor did they wholly resemble the subject's own production of English vowels.

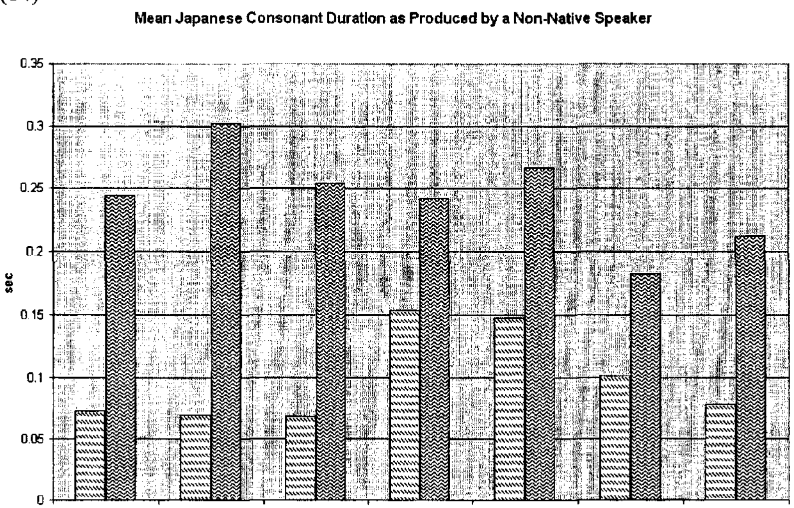
The next section aims to determine if the subject was able to make use of any linguistic knowledge gained from examination of the first recording session.

## 6. Results – Time II

### 6.1. Consonant Closure Duration

The graph in (14) below illustrates the subject's mean timing control of single and geminate consonants; Table 13 details the timing control ratios produced, as illustrated in (14).

(14)




/t/ vs. /tt/ /p/ vs. /pp/ /k/ vs. /kk/ /s/ vs. /ss/ʔ / vs. ʔ / /n/ vs. /nn//m/ vs. /mm/

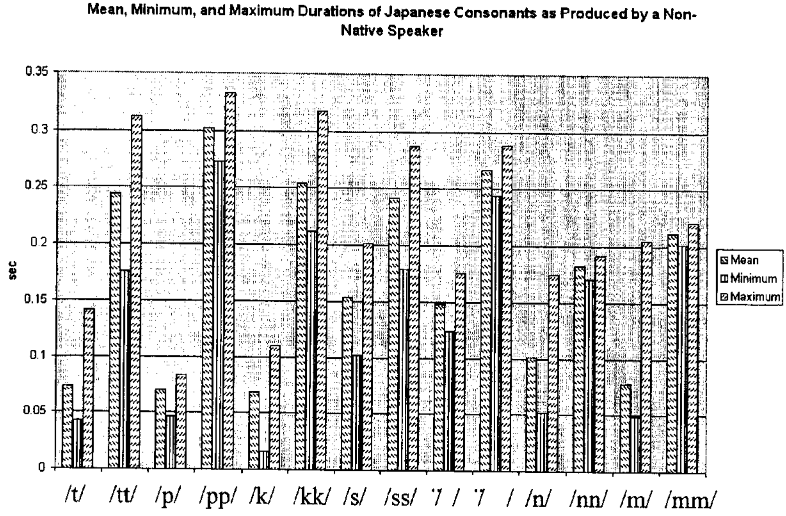
*Table 13: Mean Consonant Closure Duration Ratios*[9]

| | Single | Geminate | Ratio |
|---|---|---|---|
| /t/ vs. /tt/ | 0.073 | 0.244 | 3.34 |
| /p/ vs. /pp/ | 0.070 | 0.302 | 4.13 |
| /k/ vs. /kk/ | 0.068 | 0.245 | 3.74 |
| /s/ vs. /ss/ | 0.154 | 0.242 | 1.57 |
| 7 / vs. 7 / | 0.148 | 0.267 | 1.80 |
| /n/ vs. /nn/ | 0.101 | 0.183 | 1.81 |
| /m/ vs. /mm/ | 0.078 | 0.212 | 2.72 |

Overall, the subject produced a mean closure duration ratio of 2.59:1.0. A two-tailed *t* test revealed that geminate consonants were produced with a significantly longer closure duration than their corresponding single consonants [/t/ vs. /tt/: $t = -15.691$, $p < 0.001$; /p/ vs. /pp/: $t = -18.836$, $p < 0.001$; /k/ vs. /kk/: $t = -21.076$, $p < 0.001$; /s/ vs. /ss/: $t = -6.772$, $p < 0.001$; 7 / vs. 7 /: $t = -16.953$, $p < 0.001$; /n/ vs. /nn/: $t = -4.529$, $p < 0.001$; /m/ vs. /mm/: $t = -8.876$, $p < 0.001$], suggesting once more that the subject has indeed acquired new phonological structure enabling her to deal with the Japanese length contrast in a consistent and significant manner. Once again, performance groups coincide with natural sound classes: among the obstruents, the timing ratio patterns differently for stops (larger ratio than that reported for native speakers) than it does for continuants (smaller ratio than that reported for native speakers). Again, an inspection of the raw data reveals some variation among the tokens produced, illustrated in the graph given in (15).

[9] Recall that an additional sentence was recorded at Time II in order to obtain the /s/ vs. /s:s/ contrast, which had been omitted from Time I due to vocabulary constraints.

(15)




## 6.2. Vowel Duration

The graph in (16) illustrates the subject's mean timing control of vowel duration; Table 14 details the timing control ratios produced illustrated in (16).

(16)

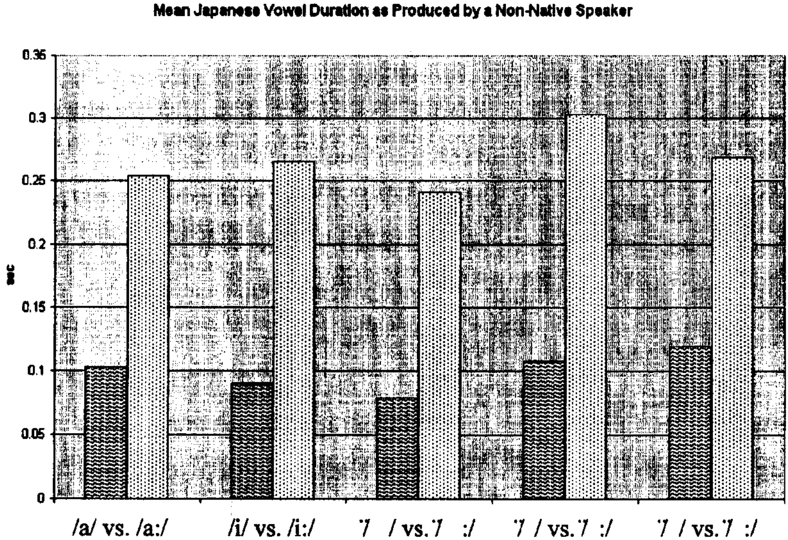


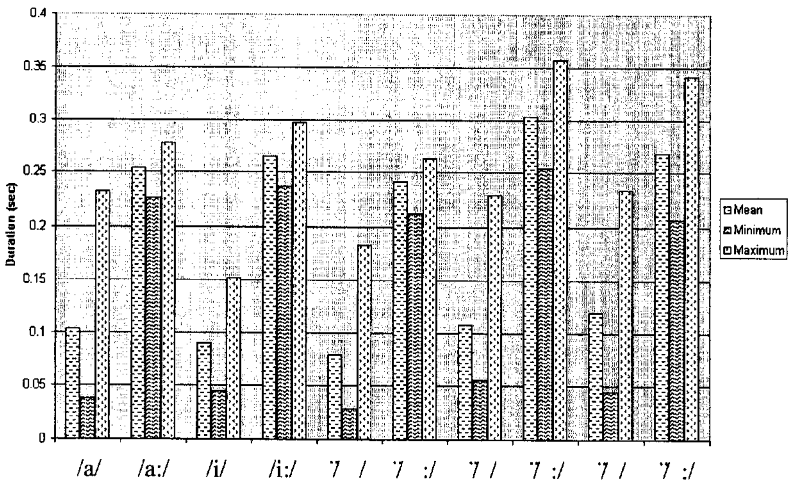

*Table 14: Mean Vowel Duration Ratios*

| | Short | Long | Ratio |
|---|---|---|---|
| /a/ vs. /a:/ | 0.103 | 0.254 | 2.47 |
| /i/ vs. /i:/ | 0.090 | 0.265 | 2.94 |
| 7 / vs. 7 :/ | 0.079 | 0.241 | 3.05 |
| 7 / vs. 7 :/ | 0.108 | 0.303 | 2.81 |
| 7 / vs. 7 :/ | 0.119 | 0.268 | 2.25 |

Overall, the subject is producing a vowel duration contrast ratio of 2.63:1.0. A two-tailed *t* test revealed that long vowel duration was significantly longer than the corresponding short vowel [/a/ vs. /a:/: $t = -10.695$, $p < 0.001$; /i/ vs. /i:/: $t = -16.203$, $p < 0.001$; 7 / vs. 7 :/: $t = -12.921$, $p < 0.001$; 7 / vs. 7 :/: $t = -12.587$, $p < 0.001$; 7 / vs. 7 :/: $t = -10.002$, $p < 0.001$]. The problematic variation that was found at Time I is still present, however; the graph in (17) illustrates this point.

(17)

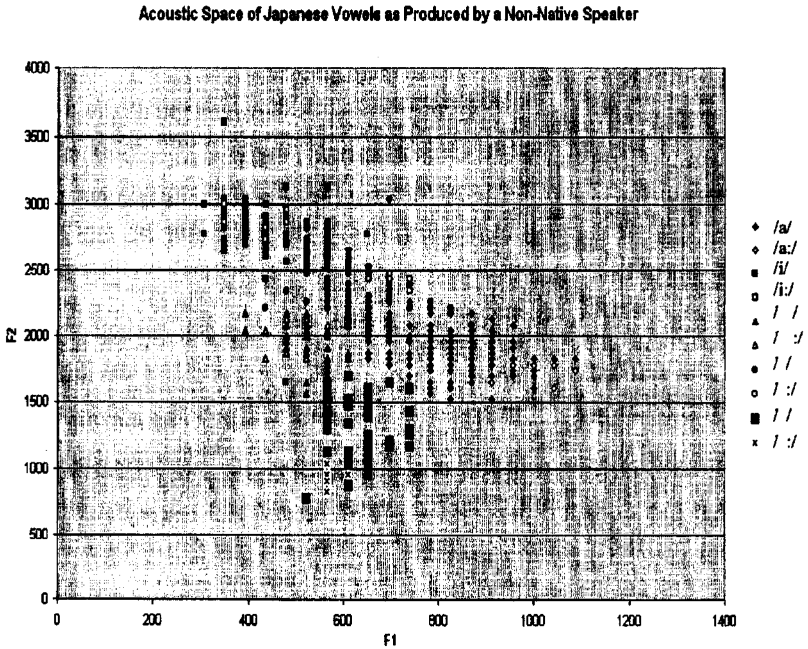
**Mean, Minimum, and Maximum Durations of Japanese Vowels as Produced by a Non-Native Speaker**




## 6.3. Vowel Quality

The scatter plot presented in (18) details the subject's vowel productions in acoustic space.

(18)




The mean F1 and F2 frequencies are given in Table 15 below, along with the mean F1 and F2 frequencies produced by a native speaker.

*Table 15: Mean F1 and F2 of Japanese Vowels*

| | F1 – Native | F1 – Non-Native | F2 – Native | F2 – Non-Native |
|---|---|---|---|---|
| /a/ | 1046 | 833 | 2075 | 1893 |
| /a:/ | 1046 | 1007 | 2075 | 1724 |
| /i/ | 354 | 454 | 2886 | 2775 |
| /i:/ | 354 | 420 | 2886 | 2876 |
| '7 / | 367 | 529 | 2060 | 1914 |
| '7 :/ | 367 | 502 | 2060 | 1980 |
| '7 / | 655 | 641 | 2209 | 2292 |
| '7 :/ | 655 | 702 | 2209 | 2427 |
| '7 / | 659 | 628 | 977 | 1244 |
| '7 :/ | 659 | 575 | 977 | 927 |

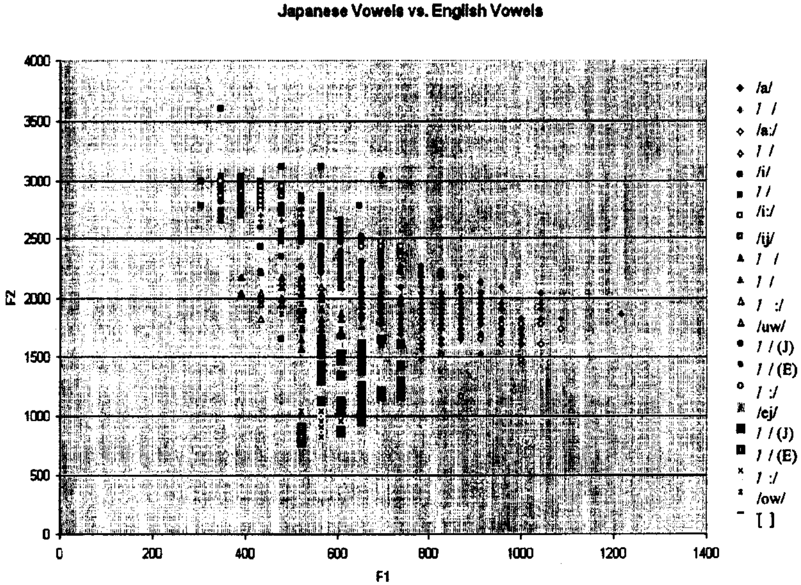
A two-tailed *t* test revealed that /a/, '7 /, and '7 / were all pronounced with a significantly different quality than their long counterpart, while /i/ and '7 / showed no significant difference between long and short counterparts. Table 16 below details the results of the statistic analysis.

*Table 16: Results of Two-Tailed t Test for Japanese Vowel Quality – Time II*

| Vowel Contrast | | t value | p value |
|---|---|---|---|
| /a/ vs. /a:/ | F1 | -4.642 | p < 0.001 |
| | F2 | 2.888 | 0.004 |
| /i/ vs. /i:/ | F1 | 1.049 | 0.297 |
| | F2 | -1.184 | 0.239 |
| '7 / vs.'7 :/ | F1 | 1.408 | 0.165 |
| | F2 | -0.946 | 0.348 |
| '7 / vs.'7 :/ | F1 | -1.949 | 0.055 |
| | F2 | -5.057 | p < 0.001 |
| '7 / vs.'7 :/ | F1 | 5.686 | p < 0.001 |
| | F2 | 7.783 | p < 0.001 |

The measurement data were plotted against the data for English vowels to check for substitutions, giving the scatter plot in (19).

(19)




A two-tailed *t* test revealed that with the exception of /i:/ and 7 :/, all Japanese vowels were pronounced with a significantly different quality than their corresponding English vowels. This suggests that the subject did not substitute English vowels for Japanese vowels in her productions, and is instead attempting to modify her vowel space so that her Japanese vowels more closely resemble those of a native speaker. The full details of statistic analysis are listed in Appendix B.

## 6.4. Comparing Time I to Time II

### 6.4.1. Consonant Closure Duration

There was no universal change in consonant closure duration from the first recording session to the second; a two-tailed *t* test revealed that many of the changes were in fact not significant. The exceptions that were found will be discussed here, the full details of the statistic analysis are listed in Appendix B.

While the change in closure duration of /t/ was not significant [$t = 1.665$, $p = 0.102$], the change in closure duration of /tt/ was found to be significantly shorter at Time II [$t = 4.202$, $p < 0.001$]. This suggests that the subject is adjusting the timing control of the geminate stop in order to obtain a more native-like ratio (3.91:1.0 at Time I vs. 3.34:1.0 at Time II). Conversely, both /p/ and /pp/ were found to be significantly shorter at Time II [/p/: t = 3.050, p = 0.008; /pp/: t = 3.147, p = 0.010], suggesting a general adjustment to the timing control of closure duration for the bilabial stop, resulting in only a slight detrimental change to the mean ratio (4.00:1.0 at Time I vs. 4.13:1.0 at Time II). A third scenario presents itself in the case of /k/ vs. /kk/ and 7 / vs. 7 /: here, the single consonant was found to be significantly shorter at Time II [/k/: t = 3.717, p < 0.001; 7 /: t = 2.878, p = 0.007], while the change in the corresponding geminate consonant was not significant [/kk/: t = 2.168, p = 0.055; 7 /: t = 2.213, p = 0.051]. In both cases, however, the mean ratio is only slightly improved by the change (/k/ vs. /kk/: 3.87:1.0 at Time I vs. 3.74:1.0 at Time II; 7 / vs. 7 /: 1.78:1.0 at Time I vs. 1.80:1.0 at Time II).

In spite of the improvement seen with respect to the /t/ vs. /tt/ contrast, these results suggest that the subject was unable to make use of information about her performance from the first recording session in the later recording. Her timing control of the geminate – single consonant contrast in most cases did not change significantly; where it did, it did not result in a great improvement to the mean ratio.

### 6.4.2. Vowel Duration

An examination of the changes in vowel duration from the first recording session to the second yielded non-uniform results, much like the analysis for consonant closure duration described in Section 6.4.1. above. A two-tailed *t* test showed that three of the five short vowels were significantly shorter at Time II [/a/: t = 3.425, p = 0.001; /i/: t = 3.030, p = 0.003; 7 /: t = 2.868, p = 0.005], whereas two of the five long vowels were found to be significantly shorter at Time II [7 :/: t = 2.628, p = 0.025; 7 :/: t = 3.146, p = 0.006] (Full details of the statistic analysis are presented in Appendix B). These changes brought about slight improvements to the mean vowel duration ratios, as illustrated in Table 17 below; the only exception to this is the increase in the 7 / vs. 7 :/ contrast ratio, which is puzzling since neither vowel in this contrastive pair was found to have significantly either increased or decreased in duration.

*Table 17: Mean Vowel Duration Ratios, Time I vs. Time II*

| | Ratio – Time I | Ratio – Time II |
|---|---|---|
| /a/ vs. /a:/ | 2.50 | 2.47 |
| /i/ vs. /i:/ | 3.22 | 2.94 |
| 7 / vs. 7 :/ | 2.67 | 3.05 |
| 7 / vs. 7 :/ | 3.08 | 2.81 |
| 7 / vs. 7 :/ | 2.29 | 2.25 |

Although slight improvement was observed across two of the five contrastive pairs of vowels, these results still suggest that the subject was unable to make use of insight on her performance gained after the first recording session. This claim is based largely on the fact that both the 7 / vs. 7 :/ and 7 / vs. 7 :/ contrasts worsened, in opposite directions, with respect to native speaker performance. If the new explicit linguistic knowledge had been available to the subject, we would have expected to see improvement across all five contrasts, not just a select few.

### 6.4.3. Vowel Quality

Turning now to an examination of changes in vowel quality from one recording session to the next, the results are again varied. A two-tailed *t* test revealed that none of the long vowels underwent a significant change in vowel quality from Time I to Time II. Among the short vowels, three did undergo a significant change, but only one did so in a manner that resulted in a more native-like vowel. The lower-than-native F1 of /a/ was significantly lower at Time II [$t = 2.438$, $p = 0.015$] and the higher-than-native F1 of /i/ was significantly higher at Time II [$t = -3.805$, $p < 0.001$]; while the lower-than-native F2 of 7 / was significantly higher (but still not native-like) at Time II [$t = -3.310$, $p = 0.001$]. The full details of the statistic analysis are presented in Appendix B.

Again, these findings suggest that the subject was not able to make use of insight on her performance gained after the first round of recordings to make her vowel production more like that of a native speaker. Only one vowel was found to have improved, two worsened, and the remainder did not undergo any significant change from one recording session to the next.

## 7. Discussion

The results of the experiment show that the subject's Japanese productions are not merely English phonemes funnelled into Japanese words and sentences, but they are not native-like either. Indeed, the data collected do show some intriguing properties, which will be addressed here, both in the context of the experiment itself, and in the context of phonological acquisition and L2 acquisition theory. Since our main research goal was to determine if explicit linguistic knowledge facilitated L2 acquisition by allowing the learner to bypass the L1 knowledge, we begin with a discussion of the results from that perspective.

In spite of some variability in the data, the single – geminate consonant contrast produced by the subject did reach significance in statistical analyses, suggesting that the subject has acquired something allowing her to distinguish reliably and consistently between single and geminate consonants. Given that this type of length contrast is not found in English, it is reasonable to assume that the subject's explicit knowledge of length contrasts in Japanese is what enabled her to build the necessary phonological structure to represent them and produce them reliably, even after only four months of classroom exposure to Japanese. When we also consider that the subject in this research outperformed one of Han's (1992) fluent subjects, who did not produce any single – geminate stop contrast at all, our conclusion grows stronger. The behaviour observed for the vowels do not provide any opposition to this conclusion, as the subject was able to produce a significant short – long vowel contrast in her Japanese productions, an outcome that was expected even under the assumption that explicit linguistic knowledge plays no role in L2 phonological acquisition.

How, then, are we to interpret the following observations: although significant, the timing ratio for consonants was never native-like, and both consonants and vowels showed variability in the tokens produced. I propose that the representation of a geminate consonant or long vowel, like all segments, has two parts: a phonological representation, and a phonetic implementation. The phonological representation would contain the abstract information that allows that segment to be distinguished phonemically from other phonemes; in the case of both geminate consonants and long vowels in Japanese, this would be a feature relating to weight. The phonetic implementation, on the other hand, would deal with the aspects pertaining to the specific articulatory details of that segment, such as timing control. The motivation for this distinction is not merely one of convenience: studies of languages that make use of a single – geminate consonant contrast found closure duration is found to be the single distinguishing cue, but the actual timing control ratios employed by the languages may differ. Thus, while Bengali and Italian both employ a 2:1 timing ratio, Turkish patterns with Japanese in making use of a 3:1 ratio (Esposito and Di Benedetto 1999; Lahiri and Hankamer 1988). Indeed, Han (1992) also distinguishes between these two levels, arguing that of the four fluent American subjects, one is evidencing a phonological difficulty and thus is unable to produce a single – geminate stop contrast, while the remaining three are evidencing a phonetic difficulty and thus are producing a contrast, but not one that resembles that produced by native speakers. It is reasonable to assume a similar situation for vowels. This distinction, then, between phonology and phonetics, allows us to account for the variability not only in the timing ratios produced but also the variability in tokens of a single segment: the subject is struggling with the phonetic representation of geminate consonants and long vowels. This may be rooted in a difficulty in perceiving the finer details of segment length, which would allow us to account for the differences in performance on stops, fricatives, and vowels. The subject may be relying on tactile cues to determine segment length: in the articulation of a stop, there is definite contact between the tongue and the articulators; a fricative is articulated with less contact; vowels are articulated with no contact at all. This accounts for the greater variability in vowel production compared with consonant production and the greater variability in fricative production compared with stop production.

Thus, explicit linguistic knowledge has been shown to be useful in building novel phonological representations, but not in developing novel phonetic implementation for the newly acquired phonological representations. This conclusion is further supported by the fact that production was not seen to improve from Time I to Time II, despite the presentation of explicit linguistic knowledge with reference to the phonetic details of Japanese length contrasts and the shortcomings of the performance at Time I: although the subject was aware of the timing ratios for consonants and vowels produced by native speakers and how her productions fell short of native-like levels at Time I, she was unable to adjust her performance so that it more closely resembled that of a native speaker at Time II. Therefore, explicit linguistic knowledge cannot be the only factor in attaining native-like proficiency in a second language.

Turning now to vowel quality, the subject was observed to produce vowels that were significantly different from her own English vowels, yet they were not like those produced by native Japanese speakers in a crucially important way: short vowels were found to differ significantly from their corresponding long vowels. Although the subject is not directly substituting English vowels in place of Japanese vowels (with the important exception of'7 / for'7 / and /uw/ for'7 :/, a substitution that is expected under Brown's (2000) model but not under Han's (1992) explicit linguistic knowledge proposal), she seems unable to make use of only one distinguishing acoustic cue in Japanese where English makes use of two. This is in sharp contrast with her ability to make use of a novel distinguishing cue in the case of geminate consonants. Furthermore, the subject was aware of the details of the Japanese vowel system, so under our explicit linguistic knowledge hypothesis she should have been able to accurately produce short vowels that were of the same vowel quality as the corresponding long vowels. We cannot, however, attribute this difficulty to the phonetic level as we did with variability of vowel duration; rather, the subject seems unable to merge what would be two vowel qualities in English into one in Japanese, and as such represents a true phonological difficulty: the subject differentiates between long and short vowels along a dimension that is not used by native speakers. Contrastively, the substitution of the rounded high back vowels of English for the unrounded high back vowels of Japanese can be accounted for as a phonetic difficulty, as roundedness is not contrastive for back vowels in either Japanese or English, therefore the subject is not motivated to build novel phonological structure and instead treats roundedness as a phonetic detail, which we have argued is not improved with explicit knowledge. Given these results, it appears that vowel quality is a domain that is not well addressed by the present explicit knowledge model and requires further research in order to fully explain the behaviour observed here. It may be the case that the problem is linked more closely to a relative difficulty of merging two phonological categories of the L1 into one single phonological category in the L2: that is, suppressing an L1 contrast that is not active in the L2. More research is required to fully examine this possibility, as the present research was aimed at examining the acquisition of new contrasts, and not the suppression of existing ones.

On the whole, the results of this research indicate that explicit linguistic knowledge is useful to the L2 learner in that it allows the learner to bypass his L1 knowledge and build novel phonological structure to accommodate the contrasts of the L2; it is not, however, useful in attaining native-like phonetic control of those contrasts. This is not the result expected under Han's (1992) proposal, which predicted that learners should be able to produce a native-like timing control ratio of geminate and single consonants if they receive explicit instruction on the nature of the timing control. Nor is this the result expected under Brown's (2000) model, which predicted poor performance on the single – geminate consonant contrasts at all stages of acquisition due to the absence of the feature for contrastive consonant length in English, thus predicting that the Japanese consonant length contrast should be impossible for the subject to acquire. This is not to say that we should dismiss Brown's (2000) model, as it does make accurate predictions about those segments in the L2 that will present a difficulty to L2 learners with different L1s. Nor do we want to dismiss Han's (1992) proposal, since it was borne out to some degree in the results of this study. Instead, the findings presented here suggest that the claims made by Brown's (2000) model about the impossibility of acquisition of certain L2 contrasts are too strong, and that it is possible for L2 learners to overcome phonological difficulties that are caused by the L1, even at very early stages of acquisition. Likewise, Han's (1992) proposal is shown to be too strong, and explicit instruction alone will not give rise to native-like performance. I would like to propose a compromise: explicit linguistic knowledge of the novel contrast will enable the learner to acquire the contrast by assisting in building the required phonological structure, but it cannot fill in the phonetic details.

If explicit linguistic knowledge is only helpful in acquisition of the phonological side of the contrast, how then is it possible for learners to acquire native-like phonetic control of the contrast? Bongaerts et. al's (2000) findings force us to acknowledge that it must be possible, yet the results presented here clearly demonstrate that explicit linguistic knowledge is not the key to success in this domain. Although more research is needed in this area, I would tentatively suggest that it is a matter of experience with the L2 that determines ability in this domain. The subject in this research performed poorly in the phonetic domain, but under this suggestion, this can be attributed to lack of experience with Japanese: at Time II, she had had only six months of classroom exposure to Japanese, which is an extremely short time in the process of language acquisition. The subjects in Han's (1992) and Bongaerts et. al's (2000) research were all advanced learners who had acquired their L2 as adults, yet only those in Bongaerts et. al (2000) were reported to perform at a level that was indistinguishable from native speakers. The crucial difference is the amount of time spent either studying the L2 or being immersed in it. Bongaerts et. al (2000) report that their star subjects had lived in the Netherlands for several years (unfortunately, no specific time frame is specified) and communicated almost exclusively in Dutch (the L2) with their families and in the work place. The subjects in Han's (1992) study, on the other hand, had spent considerably less time in a Japanese-speaking environment: one had studied Japanese intensively for two months before living in Japan for two years, one had studied Japanese in a university classroom setting for three years before spending one year studying at a Japanese university, one had spent several years living in Japan on an intermittent basis, and the fourth had studied Japanese in a classroom setting for one and a half years before spending a year living in Japan in a university student exchange program. Clearly the two groups differ with respect to the amount of experience each has had with the L2, which has impacted their ability to control the phonetic details of the language (interestingly, Han's (1992) fourth subject displayed a phonological difficulty and was unable to produce any single – geminate stop contrast in testing).

This distinction also works well when applied to the results of research on training in L2 phonological acquisition. Matthews (1997) reports on a study in which subjects whose L1 was Japanese were explicitly trained in the articulation of novel segmental contrasts in English (the L2), then tested on their perceptual abilities of these contrasts (/b/ vs. /v/, /s/ vs. 7 /, 7 / vs. /f/, and /l/ vs. 7 /). In our terms, they were given phonetic training on the contrasts. The results showed improvement, but not across all categories, and not to native-like levels, despite the subject's ability to correctly pronounce the targeted segments during training sessions. Under the analysis presented with this research, the subjects in Matthews' (1997) study were being given explicit linguistic information on articulatory (phonetic) details, which would not be helpful beyond the training session as they would not be helpful in building the phonological structure required to distinguish the novel contrasts.

Of course, the research presented should not be taken as an endpoint for investigations into this area. Of particular interest and importance to the argument drawn from this study would be an investigation into the perceptual abilities of L2 learners with explicit linguistic knowledge of the structures of the L2; the fact that the data was gathered using a reading task allows for the possibility that the subject in this particular study is responding to a visual cue in the Japanese *hiragana* script and artificially lengthening the required segments while the phonological contrast is, in fact, absent from her L2 grammar. Also of interest would be investigations into the range of timing control ratios that are deemed to be native-like by native speakers and their perceptions of differing timing ratios, both greater and lower than those of native speakers, as well as investigations into the acquisition of segmental contrasts following this same model, for although we have argued that the same mechanisms should govern both segmental and prosodic domains of phonological acquisition, this may not be the case in actual fact.

A few cautionary remarks are in order: although the results of this study suggest that explicit linguistic knowledge is useful in L2 phonological acquisition, it is not suggested that only explicit knowledge will assist the L2 learner, nor is it suggested that the presence of explicit linguistic knowledge will guarantee successful acquisition of L2 phonology. The acquisition of a second language is a complex procedure, involving many factors, such as practice and motivation, and not all factors have an equal effect on the final state of the L2 grammar. The purpose of the present research was to determine if explicit linguistic knowledge could be counted as one of these factors: does knowing about the L2 assist the learner in using it?

It should also be noted that the findings of this research are not meant to be interpreted as evidence against the Critical Period Hypothesis, nor is it suggested that all linguists will be good at learning languages. In spite of all the linguistic knowledge about Japanese at her disposal, the subject still did not achieve native-like control of the timing ratios examined, nor was she able to correct known errors in performance. L2 acquisition is still a hard and demanding task, while L1 acquisition is remarkably easy.

Thus, the present study is really just one of many steps in research into ultimate attainment in L2 acquisition. Explicit linguistic knowledge was shown to play an important role in assisting the L2 learner in acquiring novel contrasts which are distinguished by a feature absent from the L1 grammar; however, the results also indicate that explicit linguistic knowledge cannot be the only factor governing L2 phonological acquisition, as the learner was unable to achieve full native-like control of the phonetic details involved in producing the new contrast. More research is needed in order to answer the questions of why and how some adult L2 learners are able to achieve a level of phonological competence in their L2 where their performance is indistinguishable from that of a native speaker.

**Acknowledgements**
This research has benefitted greatly from the input of others. Many thanks to Michael Dobrovolsky, Martha McGinnis, Jodi Edwards, and Sheena Van Der Mark for all their comments and advice, as well as to Gisela Engels for her help with the statistical analysis. Special thanks to my supervisor John Archibald for all his patient guidance and assistance with this project. Thank you, thank you, thank you.

## Appendix A

### Test Sentences

1. Sono zassi wa takakatta desu ka.
   "Was that magazine expensive?"

2. Tiisai susiya wa eki no tonari ni arimasu.
   "The small sushi bar is beside the station."

3. Yasumi wa mikka kara youka made desu.
   "The holidays are from the third until the eighth."

4. Kinou ookii zisyo wo kaimasita.
   "Yesterday I bought a large dictionary."

5. Kuroi empitu wo ippon katte kudasai.
   "Please buy one black pencil."

6. Eeto ammari warukunai desu.
   "Well, it's not that bad."

7. Yuubinkyoku dake itte kimasu.
   "I'm only going to the post office (and I'll be right back)."

8. Yuube konnani tempura wo tabemasita.
   "Last night I ate this much tempura."

9. Daigaku no tosyokan wa benri desu nee.
"The university library is convenient, isn't it?"

10. Canada[10] no kitte ga yon mai irimasu.
"I need four Canadian stamps."

11. San zi zyuppun mae desu.
"It's ten minutes before three o'clock."

12. Kinou wa yokka zya nakatta desu. Tooka desita.
"Yesterday wasn't the fourth. It was the tenth."

13. Yasumi wa issyuukan desita.
"The holidays lasted one week."

14. Depaato[11] de furosiki wo katte kite kudasai.
"Please go buy a *furosiki* at the department store (and come right back)."

15. Sumimasen, zyaa mata.
"Excuse me, I'll see you later."

16. Kissaten de hon wo issatu yomimasita.[12]
"I read one book at the coffee shop."

---

[10] Productions of *Canada* were not included in analysis due to the obvious English origins of the word; it is likely that the subject would pronounce *Canada* in Japanese in the same manner as she would in English, thus making it unrepresentative of her pronunciation in Japanese.
[11] Although this word is borrowed from English (*department store*), it was included in the analysis because it was not deemed likely that the subject would pronounce the word as she would in English.
[12] This sentence was added at Time II in order to obtain measurements on the /s/ vs. /ss/ contrast, which had been absent in the Time I recordings due to vocabulary constraints.

## Appendix B

## Results of Statistic Analysis

*Time I*

| Consonant Duration | t value | p value |
|---|---|---|
| /t/ vs. /tt/ | -12.820 | p < 0.001 |
| /p/ vs. /pp/ | -10.649 | p < 0.001 |
| /k/ vs. /kk/ | -7.292 | 0.001 |
| '7 / vs.'7 / | -8.044 | p < 0.001 |
| /n/ vs. /nn/ | -7.640 | p < 0.001 |
| /m/ vs. /mm/ | -12.657 | p < 0.001 |

| Japanese Vowels | | t value | p value |
|---|---|---|---|
| /a/ vs. /a:/ | Duration | -9.716 | p < 0.001 |
| | F1 | -3.825 | p < 0.001 |
| | F2 | 3.464 | 0.001 |
| /i/ vs. /i:/ | Duration | -5.818 | 0.002 |
| | F1 | 2.860 | 0.016 |
| | F2 | -2.663 | 0.009 |
| '7 / vs.'7 :/ | Duration | -9.276 | p < 0.001 |
| | F1 | 1.569 | 0.123 |
| | F2 | -2.285 | 0.027 |
| '7 / vs.'7 :/ | Duration | -12.668 | p < 0.001 |
| | F1 | -1.795 | 0.077 |
| | F2 | -0.845 | 0.401 |
| '7 / vs.'7 :/ | Duration | -8.523 | p < 0.001 |
| | F1 | 3.256 | 0.004 |
| | F2 | 2.504 | 0.015 |

| **Japanese Vowels vs. English Vowels** | | t value | p value |
|---|---|---|---|
| /a/ vs.'7 / | F1 | -6.647 | $p < 0.001$ |
| | F2 | -0.509 | 0.611 |
| /a:/ vs.'7 / | F1 | 3.245 | 0.006 |
| | F2 | 1.111 | 0.287 |
| /i/ vs.'7 / | F1 | -5.309 | $p < 0.001$ |
| | F2 | 7.905 | $p < 0.001$ |
| /i:/ vs. /ij/ | F1 | -5.437 | $p < 0.001$ |
| | F2 | 0.997 | 0.362 |
| '7 / vs.'7 / | F1 | -2.114 | 0.083 |
| | F2 | 1.032 | 0.307 |
| '7 :/ vs. /uw/ | F1 | 0.923 | 0.373 |
| | F2 | -1.209 | 0.248 |
| '7 / vs.'7 / | F1 | -2.386 | 0.020 |
| | F2 | 4.161 | $p < 0.001$ |
| '7 :/ vs. /ej/ | F1 | 10.716 | $p < 0.001$ |
| | F2 | -11.939 | $p < 0.001$ |
| '7 / vs.'7 / | F1 | 2.673 | 0.010 |
| | F2 | 0.890 | 0.377 |
| '7 :/ vs. /ow/ | F1 | 2.603 | 0.022 |
| | F2 | -2.908 | 0.033 |

*Time II*

| **Consonant Duration** | t value | p value |
|---|---|---|
| /t/ vs. /tt/ | -15.691 | $p < 0.001$ |
| /p/ vs. /pp/ | -18.836 | $p < 0.001$ |
| /k/ vs. /kk/ | -21.076 | $p < 0.001$ |
| /s/ vs. /ss/ | -6.772 | $p < 0.001$ |
| '7 / vs.'7 / | -16.953 | $p < 0.001$ |
| /n/ vs. /nn/ | -4.529 | $p < 0.001$ |
| /m/ vs. /mm/ | -8.876 | $p < 0.001$ |

| **Japanese Vowels** | | t value | p value |
|---|---|---|---|
| /a/ vs. /a:/ | Duration | -10.695 | p < 0.001 |
| | F1 | -4.642 | p < 0.001 |
| | F2 | 2.888 | 0.004 |
| /i/ vs. /i:/ | Duration | -16.203 | p < 0.001 |
| | F1 | 1.049 | 0.297 |
| | F2 | -1.184 | 0.239 |
| '7 / vs.'7 :/ | Duration | -12.921 | p < 0.001 |
| | F1 | 1.408 | 0.165 |
| | F2 | -0.946 | 0.348 |
| '7 / vs.'7 :/ | Duration | -12.587 | p < 0.001 |
| | F1 | -1.949 | 0.055 |
| | F2 | -5.057 | p < 0.001 |
| '7 / vs.'7 :/ | Duration | -10.002 | p < 0.001 |
| | F1 | 5.686 | p < 0.001 |
| | F2 | 7.783 | p < 0.001 |

| **Japanese Vowels vs. English Vowels** | | t value | p value |
|---|---|---|---|
| /a/ vs.'7 / | F1 | -6.579 | p < 0.001 |
| | F2 | -0.191 | 0.849 |
| /a:/ vs.'7 / | F1 | 4.015 | 0.001 |
| | F2 | 2.262 | 0.041 |
| /i/ vs.'7 / | F1 | -3.316 | 0.001 |
| | F2 | 6.952 | p < 0.001 |
| /i:/ vs. /ij/ | F1 | -0.578 | 0.586 |
| | F2 | -0.607 | 0.552 |
| '7 / vs.'7 / | F1 | -1.600 | 0.167 |
| | F2 | 2.814 | 0.007 |
| '7 :/ vs. /uw/ | F1 | 2.101 | 0.056 |
| | F2 | -0.693 | 0.501 |
| '7 / vs.'7 / | F1 | -3.054 | 0.003 |
| | F2 | 3.251 | 0.002 |
| '7 :/ vs. /ej/ | F1 | 12.106 | p < 0.001 |
| | F2 | -12.359 | p < 0.001 |
| '7 / vs.'7 / | F1 | 3.407 | 0.001 |
| | F2 | 1.518 | 0.134 |
| '7 :/ vs. /ow/ | F1 | 2.863 | 0.013 |
| | F2 | -2.816 | 0.036 |

*Time I vs. Time II*

| **Consonant Duration** | t value | p value |
|---|---|---|
| /t/ | 1.665 | 0.102 |
| /tt/ | 4.202 | p < 0.001 |
| /p/ | 3.050 | 0.008 |
| /pp/ | 3.147 | 0.010 |
| /k/ | 3.717 | p < 0.001 |
| /kk/ | 2.168 | 0.055 |
| '7 / | 2.878 | 0.007 |
| '7 / | 2.213 | 0.051 |
| /n/ | 0.129 | 0.897 |
| /nn/ | 2.416 | 0.132 |
| /m/ | 1.782 | 0.079 |
| /mm/ | 2.591 | 0.061 |

| **Japanese Vowels** | | t value | p value |
|---|---|---|---|
| /a/ | Duration | 3.425 | 0.001 |
| | F1 | 2.438 | 0.015 |
| | F2 | -1.435 | 0.152 |
| /a:/ | Duration | 1.930 | 0.082 |
| | F1 | -0.497 | 0.630 |
| | F2 | -1.232 | 0.246 |
| /i/ | Duration | 3.030 | 0.003 |
| | F1 | -3.805 | p < 0.001 |
| | F2 | 0.323 | 0.747 |
| /i:/ | Duration | 1.856 | 0.116 |
| | F1 | -1.634 | 0.133 |
| | F2 | 1.156 | 0.275 |
| '7 / | Duration | 0.430 | 0.669 |
| | F1 | -1.810 | 0.074 |
| | F2 | -3.310 | 0.001 |
| '7 :/ | Duration | -1.882 | 0.078 |
| | F1 | -1.362 | 0.192 |
| | F2 | -0.573 | 0.575 |
| '7 / | Duration | 0.959 | 0.339 |
| | F1 | 0.929 | 0.354 |
| | F2 | 0.631 | 0.529 |

| | | | |
|---|---|---|---|
| 7 :/ | Duration | 2.628 | 0.025 |
| | F1 | 0.632 | 0.541 |
| | F2 | -2.190 | 0.053 |
| 7 / | Duration | 2.868 | 0.005 |
| | F1 | -0.287 | 0.775 |
| | F2 | -0.635 | 0.527 |
| 7 :/ | Duration | 3.146 | 0.006 |
| | F1 | 0.372 | 0.715 |
| | F2 | -0.473 | 0.643 |

## References


Archibald, J., and G. Libben, eds. 1995. *Research Perspectives on Second Language Acquisition.* Copp Clark Ltd.

Berko-Gleason, J., ed. 1997. *The Development of Language.* Allyn & Bacon.

Birdsong, D. 1999. *Second Language Acquisition and the Critical Period Hypothesis.* Lawrence Erlbaum, 1 – 22.

Bley-Vroman, R. 1989. The logical problem of foreign language learning. In S. Gass and J. Schachter, eds. *Linguistics Perspectives on Second Language Acquisition.* Cambridge University Press, 41 – 68.

Bongaerts, T., S. Mennen and F. van der Slik. 2000. Authenticity of pronunciation in naturalistic second language acquisition: the case of very advanced late learners of Dutch as a second language. *Studia Linguistica* 54 (2), 298 – 308.

Borden, G. J., K. S. Harris, and L. J. Raphael. 1994. *Speech Science Primer: Physiology, Acoustics, and Perception of Speech.* Williams & Wilkins.

Brown, C. 2000. The interrelation between speech perception and phonological acquisition from infant to adult. In J. Archibald, ed. *Second Language Acquisition and Linguistic Theory.* Blackwell, 4 – 63.

Burnham, D. K. 1986. Developmental loss of speech perception: exposure to and experience with a first language. *Applied Psycholinguistics* 7, 207 – 240.

Esposito, A., and M. G. Di Benedetto. 1999. Acoustical and perceptual study of gemination in Italian stops. *Journal of the Acoustical Society of America* 106 (4), 2051 – 2062.

Friederici, A. D., and J. M. I. Wessels. 1993. Phonotactic knowledge and its use in infant speech perception. *Perception and Psychophysics* 54, 287 – 295.

Goodsitt, J. V., J. L. Morgan, and P. K. Kuhl. 1993. Perceptual strategies in prelingual speech segmentation. *Journal of Child Language* 20, 229 – 252.

Han, M. 1994. Acoustic manifestations of mora timing in Japanese. *Journal of the Acoustical Society of America* 96 (1), 73 – 82.

Han, M. 1992. The timing control of geminate and single stop consonants in Japanese: a challenge for nonnative speakers. *Phonetica* 49, 102 – 127.
Han, M. 1962. The feature of duration in Japanese. *Study of Sounds* 10, 65 – 80.
Hohne, E. A., and P. W. Jusczyk. 1994. Two-month-old infants' sensitivity to allophonic differences. *Perception and Psychophysics* 56, 613 – 623.
Homma, Y. 1981. Durational relationship between Japanese stops and vowels. *Journal of Phonetics* 9 (3), 273 – 281.
Homma, Y. 1973. An acoustic study of Japanese vowels: their quality, pitch, amplitude, and duration. *Study of Sounds* 16, 347 - 368.
Jusczyk, P. W. 1997. *The Discovery of Spoken Language.* MIT Press, 73 – 109.
Jusczyk, P. W. 1993a. From general to language specific capacities: the WRAPSA model of how speech perception develops. *Journal of Phonetics* 21, 3 – 28.
Jusczyk, P. W., and R. N. Aslin. 1995. Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology* 29, 1 – 23.
Jusczyk, P. W., A. Cutler, and N. Redanz. 1993. Preference for the predominant stress patterns of English words. *Child Development* 64, 675 – 687.
Jusczyk, P. W., E. A. Hohne, and D. R. Mandel. 1995. Picking up regularities in the sound structure of the native language. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research.* York Press, 91 – 115.
Lahiri, A., and J. Hankamer. 1988. The timing of geminate consonants. *Journal of Phonetics* 16, 327 – 338.
Long, M. 1990. Maturational constraints on language development. *Studies in Second Language Acquisition* 12 (3), 251 – 286.
Matthews, J. 1997. The influence of pronunciation training on the perception of second language contrasts. In J. Leather and A. James, eds. *New Sounds 97.* University of Klagenfurt Press.
Newport, E. 1990. Maturational constraints on language learning. *Cognitive Science* 14, 11 – 28. Reprinted in P. Bloom, ed. 1994. *Language Acquisition: Core Readings.* MIT Press, 543 – 560.
O'Grady, W., and M. Dobrovolsky. 1996. *Contemporary Linguistic Analysis: an Introduction.* Copp Clark Ltd.
Polka, L., and J. F. Werker. 1994. Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance* 20, 421 – 435.
Port, R. F., J. Dalby, and M. O'Dell. 1987. Evidence for mora timing in Japanese. *Journal of the Acoustical Society of America* 81 (5), 1574 – 1585.

Rice, K., and P. Avery. 1995. Variability in a deterministic model of language acquisition: a theory of segmental elaboration. In J. Archibald, ed. *Phonological Acquisition and Phonological Theory*. Erlbaum, 23 – 42.
Rogers, H. 1991. *Theoretical and Practical Phonetics*. Copp Clark Pitman Ltd.
Spencer, A. 1996. *Phonology*. Blackwell.
White, L. 1989. *Universal Grammar and Second Language Acquisition*. J. Benjamins, 35 – 55.

Jennifer Mah
Department of Linguistics
University of Calgary
SS 807
2500 University Drive NW
Calgary, Alberta, CANADA
T2N 1N4
jemah@ucalgary.ca

## The Neural Substrates of Phonological Processing: An Examination of Neuroimaging Research

Jodi Edwards
University of Calgary

### 1. Introduction

Recent advances in neuroimaging technology have added a new dimension to the study of cognitive processes, enabling researchers to manipulate and observe on-line processing in the human brain. Imaging procedures, such as positron emission tomography, functional magnetic resonance imaging, and magnetoencephalography, have become widely used in the study of language processes and many studies have focused specifically on investigating the neural regions activated during phonological processing (Klein, et al., 1995; Shaywitz, et al., 1995; Price, et al., 1997; Zouridakis, et al., 1998). Phonological processing involves the encoding and analysis of the phonological attributes of a stimulus (e.g., the segmental and prosodic properties of an utterance) and can be elicited by tasks involving speech perception or phoneme discrimination (e.g., rhyme judgement tasks). Despite the use of standardized testing techniques designed to isolate phonological processes, previous attempts to localize the neural regions associated with phonological processing have produced highly variable results. The purpose of the present research is to integrate the results of a broad range of neuroimaging studies in order to identify the neural correlates of phonological processing.

In previous research, the use of positron emission tomography (PET) to determine the neural substrates of phonological processing has proved problematic. In a comparative analysis of PET imaging studies, Poeppel (1996) revealed that the majority of findings did not converge on similar regions of cortical activation, despite the use of tasks specifically designed to recruit phonological processes. In fact, across these studies, no single region was consistently implicated in the computation of phonological information, a finding that Poeppel termed a "no-overlap result" (pp. 321) for the domain of phonological processing. Although Poeppel attributed some of the variability in these results to insufficient task-control matching and problems inherent in the application of PET methodology, he also claimed that another significant contributing factor was that none of the studies were motivated by a particular theoretical framework. Consequently, Poeppel suggested that the 'no-overlap' result may actually reflect the selective activation of different aspects of phonological processing, but that, in the absence of a guiding theoretical model, the way in which the phonological tasks in these studies map onto linguistic representations during the retrieval of lexical information remains unclear.

Indefrey and Levelt (2000) proposed a model of word production that provides a theoretical foundation for the study of processes involved in lexical retrieval (see also Levelt & Indefrey, 2000). This model consists of a succession of 'core' processing components that are directly involved in the generation of words, including components specific to phonological processing. The core components each represent a characteristic processing operation, drawing clear distinctions between the linguistic functions involved in word production. For phonological processing operations, this model segregates segmental from metrical processing, with independent core components for phonological retrieval and phonological encoding. The word production model also contains a series of 'lead-in' processing components that trigger the activation of the core components at various stages in the generation of words. These lead-in processes are types of cognitive operations (e.g., visual word recognition) that are selectively activated by different kinds of cognitive tasks (e.g., word reading). As a result, these processes provide a mechanism with which to map word perception tasks onto specific linguistic representations using a model of word production.

Using their model, Indefrey and Levelt conducted a meta-analysis of a group of neuroimaging studies, focusing predominantly on PET and evoked response potential (ERP) research. They predicted that the processing components in their word production model would provide a structure for the resolution of inconsistent data from the neuroimaging research. The meta-analysis revealed a convergence of the patterns of activation observed in these studies onto a specific network of cortical regions. As a result, Indefrey and Levelt proposed that this cortical network subserves the processes involved in the generation of words and suggested that, within this network, there was functional specialization for the various processing stages of word production depicted by their model. Indefrey and Levelt concluded that their word production model provides a guiding theoretical framework for the study of language processes in neuroimaging research.

Although Indefrey and Levelt limited their meta-analysis mainly to PET and ERP research, studies using functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG) to localize the neural substrates of phonological processing have also yielded divergent results (Pugh, et al., 1996; Simos, et al., 1998; Zouridakis, et al., 1998; Poldrack, et al., 1999). The present study applies the word production model proposed by Indefrey and Levelt to a sample of fMRI and MEG studies investigating phonological processing. The first step involves decomposing the phonological tasks in these studies, using a task decomposition procedure similar to Poeppel (1996), with the assumption that these tasks involved the same processes as the phonological processing components of Indefrey and Levelt's model. If the tasks used to engage phonological processing in these studies map onto the tasks that activate the lead-in processing components outlined in the word production model, then this model will provide a theoretical framework for the integration of results from the fMRI and MEG literature. The hypothesis of the present study is that the regions of activation found in the fMRI and MEG studies will parallel the neural regions in the cortical network described by Indefrey and Levelt. This finding would suggest that current evidence for the localization of phonological processing is not inconsistent, but merely reflects activation in the various cortical regions corresponding to the different stages of phonological processing.

After a comparison of the neuroimaging techniques used to study language processes, I will describe Poeppel's review of PET methodology in previous investigations of phonological processing. Then, I will introduce Indefrey and Levelt's model of word production and compare it to alternative accounts of lexical representation, with an analysis of how each of these theories accounts for issues related to the retrieval and use of lexical information. Next, I will describe Indefrey and Levelt's meta-analysis of the data from a group of neuroimaging studies. I will then apply Indefrey and Levelt's model to the data from a sample of fMRI and MEG studies to integrate the results from these studies. An analysis of the cortical regions activated during phonological processing will follow, with a discussion about the implications of these findings for the word production model.

## 2. Comparison of Neuroimaging Techniques

### 2.1. Positron Emission Tomography (PET)

One of the first neuroimaging techniques to become widely used in the investigation of cognitive processes was positron emission tomography (PET). As a result, PET studies comprise a majority of the current empirical literature on the neural basis of cognition, in particular, research investigating the cortical regions underlying language processes (Jaeger, et al., 1996; Demb, Poldrack, & Gabrieli, 1999). Although the present study focuses on fMRI and MEG research, a description of the PET technique is useful to provide a basis for the review of previous research using this methodology and to introduce the paired subtraction paradigm used in PET research, as its application extends to the use of alternative neuroimaging techniques.

PET technology takes advantage of the fact that increased neuronal activation creates an increase in the metabolic requirements of the activated cells, such that the subsequent elevation in the regional levels of oxygen and glucose can be traced and recorded (Rugg, 1997). Monitoring changes in regional haemodynamic variables is made possible by the introduction of specific radioactive tracers into the blood stream via the injection of glucose labeled with a positron-emitting isotope (Logan, 1994). When glucose is taken up by an activated cell, positron-electron annihilation events are produced by the rapid decay of this isotope, causing positrons to radiate from the cell. As positrons are emitted from activated cells, a three-dimensional image of the cortex is generated, in which the intensity of any given neural region is proportionate to the amount of blood flow in that region (Rugg, 1997). This technique allows researchers to use haemodynamic variation as an indirect measure of relative neural activity in determining the functional neuroanatomy of various cognitive domains.

PET imaging has the advantage of producing a clear spatial representation of the cortical regions activated during a given task and, until only very recently, it was the preferred method for the assessment of cognitive function in the brain. Despite this advantage, PET imaging technology has a number of disadvantages. Although PET can provide an image with high spatial resolution, its temporal resolution is very poor. The effects of haemodynamic variation can only be observed when the increase in neural activity is sufficient enough to produce a change in the overall metabolic demands of the neuron population (Rugg, 1997). As a result, differences in the *timing* of the neuronal activity cannot be captured and the functional importance of such differences cannot be measured. Moreover, even if these effects could be observed with only a minimal increase in neural activity, because the haemodynamic response itself is delayed in relation to synaptic activity among the neurons, the observable effect would still be slower and more diffused, producing an image with poor temporal resolution.

Additional disadvantages related to PET's temporal resolution stem from the fact that changes in the haemodynamic properties of a neuron population may not necessarily produce a strong physiological signal. Because PET constructs a neural representation by capturing these changes, it requires many trials over a certain length of time to record the image. Demonet, et al., (1993) determined that the fastest rate at which PET scanning acquires an image is approximately ten seconds. These limitations on PET imaging's temporal acquisition capabilities pose a problem for the study of cognitive processes, like phonological processing, which occur in mere milliseconds.

In order to maximize the signal-to-noise ratio and reduce the effects of the limitations on data acquisition, researchers have developed a methodological strategy termed the Paired-Image Subtraction Paradigm (for a review see Fox et al., 1988; Posner et al., 1988; Friston et al., 1993). This paradigm requires the presentation of two minimally different tasks: a control task and an experimental task. The experimental task is a more complex extension of the control task, assumed to engage all of the cognitive processes activated by the control task, plus an additional process of interest. The control image is then subtracted from the experimental image and the resulting area of activation is argued to be the neural substrate for the process by which the two tasks differ. Although this paradigm creates an adequate signal-to-noise ratio, in order to obtain enough statistical power to interpret subtle changes in this signal, it is necessary for the data to be averaged across participants (Fox et al., 1988). One limitation with this procedure is that pooling the data requires each image to be compared with a 'standardized' image. These comparisons fail to take into account individual differences in cortical structure, as averaging the images across subjects decreases the quality of their spatial resolution. As a result, the paired subtraction paradigm precludes the study of individual differences observed in the patterns of activation associated with each task.

Another important drawback of PET methodology is that there is a limit to the number of times a participant can undergo this testing procedure (Rugg, 1997). Positrons are not emitted very frequently from cells and, as such, numerous scans may be required from each participant in order to obtain interpretable results. For PET, the need for multiple samples is particularly problematic, as it is a relatively invasive procedure, and each participant can only undergo a restricted number of scans in order to maintain safe levels of radiation exposure. Due to its drawbacks, the use of PET imaging has declined for the study of cognitive processes, however, another haemodynamic method that has become widely used to image both the structure and the function of the brain is functional magnetic resonance imaging (fMRI). Since the majority of the studies in Indefrey and Levelt's (2000) meta-analysis involved the use of PET scanning, the present research focuses, in part, on fMRI studies investigating the structure and function of the cortical regions involved in phonological processing.

### 2.2. Functional Magnetic Resonance Imaging (fMRI)

One type of fMRI procedure that enables researchers to map cortical structure and function is the 'blood oxygenation level dependent' (BOLD) method (Ogawa et al., 1990, 1992). The BOLD method involves indexing the variability in the levels of blood oxygenation in the cortex, such that the effect of oxygen concentration on various types of haemoglobin in the blood serves as a measure of neural activity. When the metabolic demands placed on cortical cells are increased due to greater neuronal activation, the corresponding increase in the levels of oxygen in the blood supply exceeds what is necessary to meet these metabolic demands. This excess creates regions of richly oxygenated cells. Oxygenation levels affect the magnetic properties of haemoglobin agents in the blood and, when placed within a magnetic field, deoxyhaemoglobin has a greater magnetic susceptibility than oxyhaemoglobin (Rugg, 1997). As the concentration of deoxyhaemoglobin decreases, a signal that is sensitive to these changes will reflect the ratio of deoxy- to oxyhaemoglobin within the blood, providing information about the density of activated neurons in the region and the intensity of their activation. This information can be used to construct an image of the cortex in response to various stimulus types.

fMRI technology presents a number of improvements over PET imaging. Most notably, because fMRI is a non-invasive procedure, there is no limit on the number of scans that can be obtained from a single participant (Rugg, 1997). This increases the reliability and power of within-subject effects and allows for the analysis of individual differences in the patterns of cortical activation. The data acquisition abilities of fMRI technology are also more advanced, increasing the speed of image acquisition from approximately ten seconds per image for PET scanning, to only three seconds per image using fMRI. fMRI also offers the advantage of high spatial resolution in its images. Although PET produces an image with high spatial quality, it records strictly functional responses, whereas fMRI produces a fine-grained representation of the anatomical structure of the cortex. The benefits associated with fMRI technology have extended its application to the study of many cognitive domains and it is increasingly employed to isolate the neural correlates of language function.

Despite its broadening application as a more advanced imaging technique, fMRI also suffers from some significant disadvantages. One drawback, inherent to the BOLD methodology, is that the BOLD signal undergoes degradation from cells in certain cortical regions that are sensitive to variation in magnetic properties other than those changes produced by blood oxygenation (Rugg, 1997). As a result, fMRI is not equally sensitive to the distribution of neuronal activation across all regions of the cortex. Another factor that contributes to the degradation of the signal quality is head movements. This is problematic for the experimental use of most imaging techniques, including fMRI, as it limits the kinds of testing procedures used in the design (e.g., even the minimal movements involved in speech production tasks introduce artifacts into the data) and requires participants to remain very still throughout the experiment.

Additionally, although fMRI offers an improvement over PET in the speed of image acquisition, fMRI images still have very poor temporal resolution. Despite their spatial capabilities, all imaging techniques that rely on haemodynamic variation to provide an indirect indication of neural activity are severely restricted in their ability to index the time course of this activity. An alternative to the use of haemodynamic imaging methods is the application of electrophysiological techniques that have the ability to measure the neural correlates of processing in real time. One such technique, that measures electromagnetic responses in the cortex, is magnetoencephalography (MEG). MEG has yet to become widely used in the study of cognitive function and, as such, very few studies in Indefrey and Levelt's meta-analysis involve the use of this technique. The present study involves an analysis of MEG research, in addition to studies using fMRI, to provide an overview of current evidence for the localization of the neural substrates of phonological processing from multiple research perspectives.

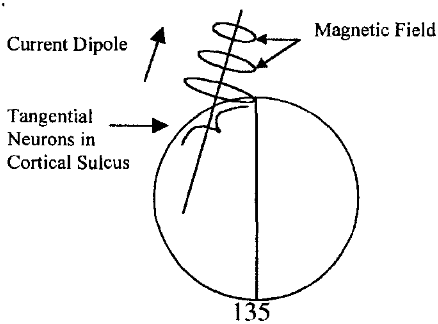
### 2.3. Magnetoencephalography (MEG)

When regional neuronal activity exceeds the threshold of background activity levels, electrical currents in the population of activated neurons produce a detectable electromagnetic field. The magnetic flux associated with this electromagnetic field can be measured using magnetoencephalography (MEG) (for a review see Hämäläinen et al., 1993). MEG takes advantage of the fact that neurons, when activated, undergo a change in polarity, becoming either polarized or depolarized (Rugg, 1997). This process produces electrical currents in the neurons that generate magnetic fields. The strength of these magnetic fields is measured in proportion to the electroencephalographic waves emitted by neuron populations in the brain during cortical activation. The measured responses are represented as waveforms of varying frequencies that can be quantitatively analyzed to produce topographic maps of the intensity and direction of the brain's magnetic fields, which may be used to infer the localizations of current dipoles in the cortex.

The greatest benefit associated with MEG technology is the high temporal resolution of the images it generates. The ability of MEG to track the correlates of on-line cognitive processing allows researchers to establish an initiation point for stimulus discrimination. When coupled with well-established behavioural results (e.g., reaction time studies), these studies of the temporal properties of neural activity can only serve to enhance our understanding of information processing. Another advantage of MEG is that it provides another non-invasive alternative to the study of cortical activity. Just as with fMRI, this allows for the application of an experimental design with numerous conditions, as each participant may undergo multiple recording sessions.

One important limitation for MEG, however, is that the magnetic recordings are sensitive only to fields generated by current dipoles in the neurons orientated *tangentially* to the scalp (see Figure 1) (Rugg, 1997).

(1) Figure 1.



As a result, the majority of MEG recordings capture activation localized exclusively to cortical sulci and reflect a minimal contribution from activity in the cortical gyri. Since a significant portion of the cortex is comprised of gyral neurons, this restricts the usefulness of MEG in many contexts. However, Phillips et al. (1995) argued that much of the activation related to auditory language phenomena is localized to the lateral sulcus and, as a result, this type of language processing can be easily measured using MEG.

In summary, neuroimaging techniques, such as PET, fMRI and MEG, provide a valuable tool for the analysis of on-line processing in the human brain. These techniques provide researchers with the potential to make more direct observations of various cortical language functions that previously could only be inferred indirectly from lesion studies or developmental disorders. With new technological capabilities, however, comes a new set of methodological issues. The application of neuroimaging methodology to the study of language processing has produced variable results and a review of previous neuroimaging research demonstrates that current evidence for the cortical localization of phonological processing remains unreliable.

## 3. Previous Research

Previous attempts to isolate the neural correlates of phonological processing using neuroimaging technology have yielded divergent results. Poeppel (1996) examined some of these results in a review of PET studies investigating the neural basis of phonological processing and revealed that many questions remain about the functional neuroanatomy of phonological processing.

### 3.1 Poeppel (1996)

Poeppel (1996) conducted a critical review of five PET studies designed to locate the neural correlates of phonological processing (see Table 1). Although all of the studies used PET methodology combined with experimental task paradigms designed to recruit phonological processing, Poeppel observed that each study reported different, ***non-overlapping***, regions of cortical activation. In order to determine the cause of such varying results, Poeppel conducted a decomposition of the experimental tasks employed in these studies. For the task decomposition, Poppel analyzed each of the studies, focusing on how the tasks of each experiment were designed to engage phonological processing specifically. The following sections summarize the conclusions drawn by the authors of each of these studies about the neural basis of phonological processes and Poeppel's analysis of these conclusions.

(2) Table 1.

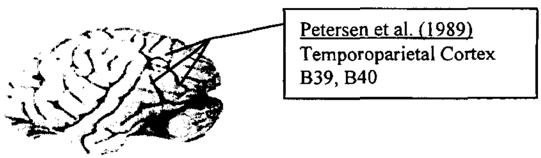
| Study | Task Type | Stimulus Type | Modality | Area of Activation |
|---|---|---|---|---|
| 1. Petersen et al. (1989) | Rhyme Judgement Task | Word Pairs | Visual | •Temporoparietal cortex |
| 2. Zatorre et al. (1992) | Rhyme Judgement Task | Word and Non-Word Pairs (CVCs) | Auditory | •Left posterior temporal and parietal regions<br>•Broca's area |
| 3. Sergent et al. (1992) | Rhyme Judgement Task | Single Letters (Consonants) | Visual | •Prefrontal areas |
| 4. Demonet et al. (1992) | Phoneme-Monitoring Task | Non-Words | Auditory | •Superior, middle and inferior temporal gyri<br>•Broca's area |
| 5. Paulesu et al. (1993) | Rhyme Judgement Task | Single Letters (Consonants) | Visual | •Superior temporal gyrus<br>•Broca's area<br>•Temporoparietal cortex |

### 3.1.1. Petersen et al. (1989)

Petersen et al. (1989) were among the first to investigate language processes using PET imaging methodology, and phonological processing was only one of the experimental processes in this large-scale study designed to investigate the neural basis of language. In order to recruit phonological processing specifically, Petersen et al. used a rhyme judgement task in which participants were visually presented with a pair of words, one above and one below a fixation point, and were asked to indicate whether the word pair rhymed. The stimuli consisted of either visually similar, rhyming, word pairs (e.g., dog-bog) or visually dissimilar, nonrhyming, word pairs (e.g., dog-cat). As a control condition, participants were presented with the same stimuli, but were not required to make a response.

A number of regions displayed a significant increase in regional blood flow during the experimental condition. Once Petersen et al. subtracted the control condition from the test condition, the resulting pattern of cortical activation was localized to left temporoparietal regions (see Figure 2). As a result, Petersen et al. argued that the left temporoparietal cortex mediates phonological processing or encoding.

(3) Figure 2.



### 3.1.2. Zatorre et al. (1992)

Zatorre et al. (1992) investigated which cortical regions were sensitive to speech stimuli by auditorily presenting word and pseudo-word pairs consisting of consonant-vowel-consonant (CVC) strings. In the experimental condition, designed to isolate phonological processing, participants had to judge whether the final consonants in the CVC pair were the same (e.g., bag-big) or different (e.g., tig-lat). As a control, they conducted a passive-listening condition, in which participants were presented with the stimuli from the experimental condition, but were not required to generate a response. Poeppel argued that this design did not allow Zatorre et al. to isolate phonological processing, as presumably participants would still make a same-different judgement, regardless of whether or not they articulated that judgement. However, Zatorre et al. argued that this design would activate phonological processing and used the paired subtraction paradigm to determine the localization of activation for the phonological condition.

Their results for the phonological condition indicated a significant increase in the regional blood flow to left-lateralized regions, including Broca's area and temporo-parietal cortex (see Figure 3). Zatorre et al. suggested that this pattern of activation implicates a cortical network underlying phonological processing involving left posterior temporal and parietal regions and Broca's area.

(4) Figure 3.

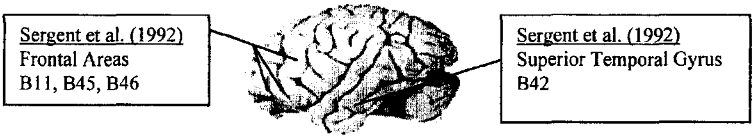

### 3.1.3. Sergent et al. (1992)

In an attempt to replicate the results of Petersen et al. (1989) or Zatorre et al. (1992), Sergent et al. (1992) also conducted a study using a rhyme judgement task, however, they maintained that the only way to isolate one specific aspect of linguistic processing, such as phonological processing, was to use stimuli that were not subject to interference from other linguistic information, such as semantic, or orthographic, codes. Accordingly, they used letter stimuli in a letter-sound task, where participants were required to make a rhyme judgement between individual, visually presented letters and an auditorily presented speech sound (e.g., does <e> rhyme with [iy]). Sergent et al. argued that this task would isolate the visual and phonological codes of the letter without generating interference from other types of linguistic processing.

Using the paired subtraction paradigm, Sergent et al. devised two control conditions for subtraction from the experimental condition. To subtract out activation from semantic codes, they used an object task, involving a forced-choice categorization about the 'living' status of line-drawings that depicted either living (e.g., a person) or non-living objects (e.g., a house). In order to control for interference from orthographic codes, the second control condition was a letter-spatial task, involving visually asymmetrical consonants that either rhymed (e.g., <b-c-d-g-p-z>) or did not rhyme (e.g., <f-j-k-l-n-r>) with the auditorily presented speech sound (e.g., [iy]). These letters were either oriented normally or in an upright mirror-reversed position and participants were required to make a forced-choice orientation judgement.

Importantly, the results of this replication study did not converge with those found by either Zatorre et al. (1992) or Petersen et al. (1989). Instead, after the subtraction of both control conditions, Sergent et al. observed significant activation in three prefrontal areas of the cortex (see Figure 4). Sergent et al. attributed this pattern of activation to the involvement of regions specialized for the programming of articulatory gestures, yet Poeppel argued that this explanation was insufficient, considering the fact that neither the experimental or control condition required any explicit articulation. It is apparent that this attempt to replicate previous neuroanatomical findings for the domain of phonological processing only contributed to the discrepancy among PET results.

(5) Figure 4.



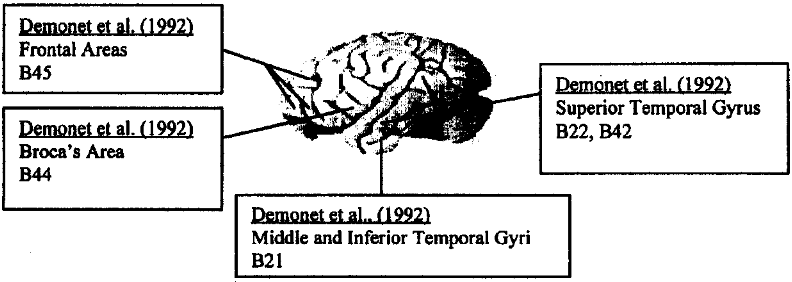

### 3.1.4. Demonet et al. (1992)

Demonet et al. (1992) investigated whether or not separate cortical regions were implicated in semantic and phonological processing. In order to isolate phonological processing, they employed a "phoneme-monitoring" task using auditorily presented multisyllabic nonwords. Participants were required to monitor the stimulus for the presence of a specific sequence of speech sounds (e.g., a [d] followed by a [b]). A non-speech sound control condition involved groups of pure tones, in which the participants were required to determine if there was a rising pitch in the third tone.

The results of the subtraction showed an increase in the regional blood flow to the superior, middle and inferior temporal gyri and Broca's area (see

Figure 5). The authors attributed the most significance to activation in regions of the superior temporal cortex. They concluded that phonological processing activates auditory association cortex in the left superior temporal gyrus and the anterior part of Wernicke's area, but suggested that Broca's area may also play a role in phonological processing.

(6) Figure 5.

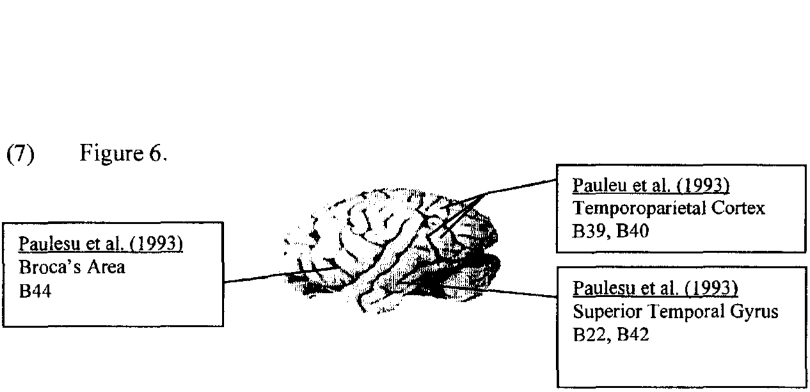



### 3.1.5. Paulesu et al. (1993)

Although the research of Paulesu et al. (1993) focused on the verbal components of working memory, they also made specific claims about the neural substrates of phonological processing. In their memory model, based on Baddeley (1992), phonological processing is a combination of phonological encoding and rehearsal operations and, as such, their experiment employed two language conditions designed to capture these processes. The first verbal task consisted of a sequence of six visually presented consonants that were phonologically dissimilar (e.g., <k-l-m-p-q-s>). In a dual-task paradigm, participants were required to maintain this string of consonants by subvocal rehearsal to activate both phonological encoding and subvocal rehearsal processes, while monitoring for a visually presented probe stimulus (e.g., <f>). The second was a rhyme judgement task for visually presented letters that was intended to activate only phonological encoding (e.g., does the target rhyme with /iy/). The results from these two language tasks were combined and then compared to a control condition, in order to separate the memory component from the verbal component.

The results of the verbal component indicated patterns of activation in areas including Broca's area, superior temporal gyrus, the supramarginal gyrus of the temporo-parietal cortex and the insulae in primary auditory cortex (see Figure 6). Paulesu et al. concluded that phonological encoding can be localized to the supramarginal gyrus of the left hemisphere and that areas in the superior temporal gyrus and frontal cortex may also be involved in phonological processing.

(7) Figure 6.




The results of the Paulesu et al. study are particularly striking in comparison to the experiment conducted by Sergent et al. (1992). These two experiments used exactly the same type of task to recruit phonological processing (e.g., a rhyme judgment task in the visual modality) with single consonant stimuli, but the localization results are dramatically different. Where Sergent et al. reported activation in prefrontal areas for phonological processing, Paulesu et al. concluded that phonological processing is localized to regions in the temporal and temporoparietal cortex and Broca's area. This comparison illustrates the high degree of variability among studies using PET methodology to identify the neural substrates of phonological processing.

**3.2. Summary**

Poeppel's comparison of PET results revealed that, despite our ability to combine cognitive testing with neuroimaging technology, current data for the cortical localization of phonological processing is highly variable. Poeppel (1996:321) termed this a 'no-overlap' result for the domain of phonological processing and concluded that it is premature to infer which cortical areas mediate phonological processing from PET research. In his analysis, Poeppel argued that problems with the experimental designs in these studies and in the application of PET methodology may account for some of the variability in the PET data. However, he also observed that none of the experiments in his review were motivated by a particular theoretical model. As a result, Poppel argued that the lack of reference to a specific theoretical framework may account for a *significant* amount of the variability in these results. Indeed, Poeppel proposed that a possible explanation for the no-overlap result could be that the tasks in these studies engaged different aspects of phonological processing, but that, without a theoretical model detailing the various aspects of phonological processing, the results appear inconsistent.

The benefit of contact with a theory of phonological processing for these experiments is that theoretical models provide independent evidence for what processes are involved in language function and how these processes interact.

Phonological processing does not involve a single operation, but a complex combination of operations related to the analysis of phonological information. As a result, phonological tasks, although they may appear similar, may actually be activating very different types of phonological processes. For instance, a task that involves word reading may activate processes related to the retrieval of phonological information for specific word forms from the lexicon. By contrast, a task that involves the identification of non-words, which have no existing lexical form, may activate processes associated with the conversion of graphemes to phonemes. The patterns of cortical activation for each of these types of phonological processes may be localized to different regions of the cortex. As a result, reference to a theoretical model of language processing may account for apparently divergent neuroimaging results.

## 4. Theoretical Background

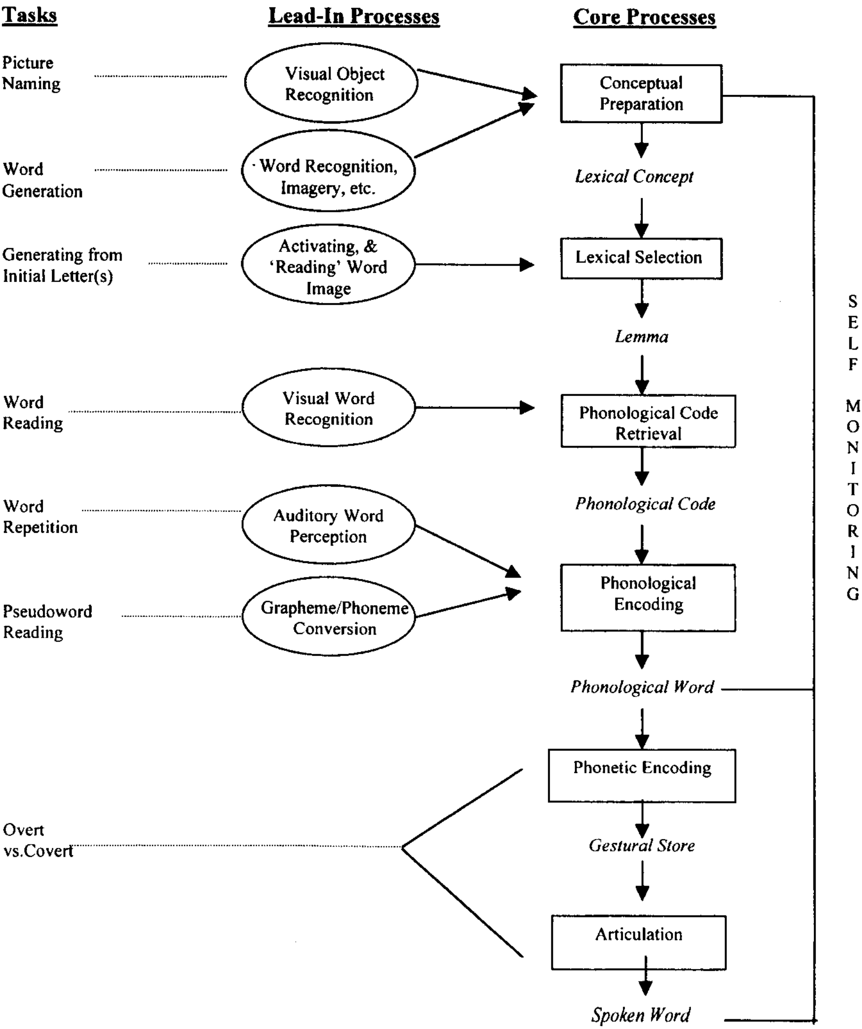
There is a large body of literature in linguistics devoted to lexical and phonological theory and many of these theories provide detailed models of the functions involved speech production and perception, including the retrieval and use of phonological information. Indefrey and Levelt (2000) proposed a theory of word production involving a staged succession of processing operations activated during the retrieval of lexical information. In the following sections, I will introduce and describe Indefrey and Levelt's word production model and then compare it to alternative accounts of lexical representation from the linguistics literature.

### 4.1. The Word Production Model

The model of word production proposed by Indefrey and Levelt (2000) depicts a succession of stages involved in the generation of a phonological word from a conceptual base (see Figure 6). This model, based on Levelt's (1989) theory of word production, involves a series of processing components that function in staged succession throughout the generation of speech. This succession contains a group of 'core' processing components (e.g., phonological code retrieval) and a corresponding set of 'lead-in' processes (e.g., visual word recognition). The core components are directly involved in the generation of words, whereas the lead-in processes enter into the word production model at various stages in the succession of the core processing components and indirectly influence the process of word production. The lead-in processes are selectively activated by different kinds of cognitive tasks (e.g., word reading) and, as such, provide a mechanism with which to map these tasks onto linguistic representations.

(8) Figure 6.

The processing stages involved in the generation of words (core processes) and the processes activated by various experimental word production tasks (lead-in processes), as depicted by Levelt & Indefrey (2000: 80).



In this model, the core processing components each have a characteristic linguistic input. Each core component selects its characteristic input and performs a computation on that input, which generates the linguistic output for that component. This linguistic output then functions as the input for the next core processing component. Each lead-in process involves a specific type of on-line perceptual processing (e.g., visual word recognition) that enters the word production mechanism at a specific stage in the sequence of core processes and triggers the activation of the core process at that stage. The type of on-line perceptual task determines which lead-in process is activated and at which level this process enters the word production mechanism. If a lead-in process activates a core process in the middle of the word production mechanism, only the processing components subsequent to this component will, in turn, be activated. Importantly, it is the differentiation between the core and lead-in processes in this model that enables studies of word *perception* to be analyzed in terms of a word *production* model. Many studies have used perceptual task designs to investigate specific aspects of language processing, such as phonological processing, and the word production model has core components designed to account for the processing of phonological information.

### 4.1.1. Phonological Processing

The core processing components of the word production model include two processing units specialized for phonological processing. The first, phonological code retrieval, selects a 'lemma' as its characteristic input (e.g., the noun lemma 'CAT + plural'). Lemmas are syntactic words characterized by a syntactic frame. The lexical concepts for verbs and function words each have a lemma whose syntactic frame specifies how the information in the lexical concept (e.g., semantic information) should be mapped onto syntactic category information (e.g., transitive verb). The syntactic frames of the selected lemmas (e.g., verbs and nouns) undergo the late insertion of phonological information as they combine to incrementally generate the syntactic pattern of the phrase, known as the 'surface structure'. This process of syntactic construction begins as soon as the lemma is retrieved (Indefrey & Levelt, 2000: 847). In order to generate the appropriate output, the phonological code retrieval component must retrieve the phonological codes from the selected lemma. The selection of the lemma triggers the activation of the phonological code for each of its morphemes (e.g., 'CAT + plural' activates /kæt/ + /-z/). The newly generated phonological code is the output of the phonological code retrieval processing component and, as such, becomes the input for a second phonological processing component, phonological encoding.

In the phonological encoding component, the primary function of the phonological code is to provide information for the generation of syllabic structure. The process of syllabification is dependent, in part, on inflectional morphology. When the lemma activates the phonological code, the inflectional information contained in the lemma indirectly influences the syllabification process. The domain of syllabification is the phonological word and as these words are syllabified, they become the output of the phonological encoding processing component. In subsequent syntactic processes, phonological words may be incrementally syllabified into connected speech and, as the surface structure expands, these words may combine to form larger units, such as phonological phrases.

The 'lead-in' process that engages the word production mechanism at the level of phonological code retrieval is visual word recognition. Visual word recognition is a type of perceptual processing that is activated during any number of word reading tasks. The lead-in processes that trigger activation at the level of phonological encoding are auditory word perception and grapheme/phoneme conversion. The process of auditory word perception occurs during tasks of word repetition and grapheme/phoneme conversion processes are utilized for pseudoword reading.

In summary, Indefrey and Levelt's theory of word production is based on a complex system of processing components that allow cognitive testing techniques to be mapped directly onto the linguistic representations involved in lexical retrieval. There are a number of issues that a theory of lexical representation must account for in order to characterize the processes involved in the retrieval and use of lexical information. Some of these issues relate to lexical access in general, such as modularity and the specification of features in the lexicon, and some relate specifically to the processing of phonological information, such as the time course for the insertion of phonological information. Indefrey & Levelt integrate these issues into their model of word production and, in the next section, I will compare their model to alternative accounts of lexical representation from the linguistics literature.

### 4.2. Theories of Lexical Representation

Many linguistic theories of lexical representation propose modular systems of representation for the linguistic functions involved in accessing information from the lexicon. In modular models of lexical retrieval, the individual components of the model each conduct linguistic operations independent of one another, usually within a serial processing framework. For example, in Indefrey and Levelt's model of word production, each processing component, or module, has a characteristic input, on which it conducts a specific processing operation in order to generate the input for the next module in the system. Current morphological theories also advocate a modular approach to lexical representation, beginning with how information is represented and processed within the lexicon.

Morphological theories of lexical representation offer a number of alternative suggestions about the nature of the information in the lexicon. In the majority of models, lexical items are comprised of a set of stems and a set of affixes that are stored in the lexicon and projected into the syntax during the generation of word forms. However, differences in these models emerge concerning exactly what type of information these lexical items are specified for. Lieber (1992) proposed a model with affixes fully specified for feature information (e.g., number, lexical category, tense). By this account, the items in the lexicon provide enough information to generate the correct structures for the syntax. By contrast, in Indefrey and Levelt's model of word production, lexical items are not specified for feature information and, as such, combine with both inflectional and derivational information prior to leaving the lexicon. Similarly, in Wunderlich's model of inflectional morphology, termed 'Minimalist Morphology' (1997), lexical items are maximally underspecified for feature information. Feature underspecification minimizes the amount of information to be stored in the lexicon, but also usually requires that lexical items combine with morphological information prior to entering the syntax in order to generate the appropriate structure for syntactic processing.

Some morphological models support the notion that the addition of derivational and inflectional information is restricted to the lexicon, while others suggest that the combination of elements occurs outside the lexicon. In 'lexicalist' approaches to morphology, the realization of elements in the syntax is governed by the lexicon. For example, in their lexicalist models, both Indefrey and Levelt, and Wunderlich, proposed that processes combining roots with morphological information occur strictly within the lexicon, prior to entering the syntax. In Wunderlich's model, both morphosyntactic and morphophonological features are added into the lexicon via morphological operations and subsequent syntactic mechanisms (e.g., agreement and case checking) are based on the availability of this information in the word forms. Although Indefrey and Levelt advocate the early insertion of morphosyntactic information, the addition of morphophonological information occurs later, in the phonological processing components of their model.

Like these lexicalist models, Lieber (1992) also proposed that lexical items combine within the lexicon to form words that undergo syntactic operations. By this account, the insertion of morphophonological information occurs early in the system. However, for Lieber, this morphological information combines with lexical roots via syntactic processes, initiating a departure from the traditional lexicalist view of morphology and lexical representation.

Halle and Marantz (1993) proposed an alternative to the lexicalist approach with their 'Distributed Morphology' model that focuses on the syntax, rather than the lexicon. For Halle and Marantz, morphological operations are distributed among several different components in the grammar and both derivational and inflectional morphology are inserted after the lexicon. In this account, morphosyntactic features are added in the syntax via syntactic operations (e.g., merge and movement operations: Chomsky, 1993), whereas, like Indefrey and Levelt, the addition of morphophonological features occurs later by vocabulary insertion in the morphology/phonology components of the system.

Another model that departs from lexicalist views is the 'a-morphous', or affixless theory of morphology (Anderson, 1992). This view is based on the notion that the lexicon does not contain affixes, but is composed of roots and word formation rules only. In this approach, there is a separation of derivational and inflectional information, where derivational information is added to the roots in the lexicon and inflectional information is inserted in the syntax via word formation rules. Although Anderson's model supports the early insertion of derivational morphophonology, again there is a division in this model, as, unlike Indefrey and Levelt and Halle and Marantz, the addition of *inflectional* morphophonology occurs later in the system.

In summary, the word production model proposed by Indefrey and Levelt combines aspects of Wunderlich's 'Minimalist Morphology' approach and the 'Distributed Morphology' model of Halle and Marantz. Indefrey and Levelt's theory is similar to Wunderlich, as both offer models in which word roots combine with derivational and inflectional information in the lexicon. However, these models diverge on the issue of phonological insertion. Wunderlich proposes that all morphological information, including morphophonological information is inserted in the lexicon, but Indefrey and Levelt make the assumption that morphophonological information is inserted late, or outside of the lexicon. On this issue, Indefrey and Levelt's model of word production is more like the 'Distributed Morphology' model. Halle and Marantz also maintain that morphophonological information is inserted late, but, unlike Indefrey and Levelt, the combination of lexical items with morphological information occurs in the syntax and not in the lexicon. The model of word production proposed by Indefrey and Levelt provides a detailed account of the processes involved in the retrieval of lexical information. If, for a given processing component, there is an underlying region of cortical function, then this region should be activated by any task involving the processes for that component. To test this claim, Indefrey and Levelt conducted a meta-analysis of a group of neuroimaging studies in an attempt to identify the neural substrates of the different processing components described in their model.

### 4.3. The Meta-Analysis

Indefrey and Levelt (2000) conducted a comprehensive meta-analysis of cerebral activation results from fifty-eight brain imaging experiments. Of these studies, the majority used PET or ERP methodology and only five studies employed fMRI or MEG techniques. In order to combine data from methods with varying resolution capabilities, Indefrey and Levelt used a double reference system to determine the significance of an activation. This system involved the comparison of localizations at both gross and fine-grained divisions of the cortex. They accepted the convergence of evidence onto a certain region only if the chance level given by a binomial distribution was below 10%. This strict threshold for significance meant they considered the agreement of reports of activation in a certain region to be coincidental unless their statistical analysis proved that there was a less than 10% chance that the convergence of these results would occur. The results of their statistical analysis showed a significant pattern of convergence onto a specific network of cortical structures. This finding revealed that Indefrey and Levelt were successfully able to determine reliable localizations for the different processes in their model.

Indefrey and Levelt made specific claims about the cortical regions activated by the core processing components of the word production model. They proposed that a left-lateralized word production network, involving both cortical and subcortical structures, subserves the core processes of word production up to and including phonological encoding. This network consists of the posterior inferior frontal gyrus (Broca's area), the mid-superior and middle temporal gyri, the posterior superior and middle temporal gyri (Wernicke's area), and the left thalamus. For the lead-in processing components, Indefrey and Levelt implicated cortical regions within the general word production neural network and suggested that different patterns of activation within this network reflect the activation of different lead-in processes.

The phonological components of the word production model displayed characteristic patterns of cortical activation. For phonological code retrieval the lead-in process is visual word recognition. In these studies, tasks of word reading that actively engaged the process of visual word recognition generated activation in the left posterior superior and middle temporal gyri (e.g., Wernicke's area) and the left thalamus. As a result, Indefrey and Levelt identified these regions as the neural substrates of phonological code retrieval. The lead-in processes for phonological encoding are auditory word recognition and grapheme/phoneme conversion. Indefrey and Levelt determined that no specific pattern of activation corresponded perfectly to these processes, but the trend in studies using word repetition and pseudo-word reading tasks to engage these processes was activation in the left mid superior temporal gyrus and the left inferior frontal gyrus (e.g., Broca's area). Indefrey and Levelt proposed that these are likely the regions underlying the core process of phonological encoding.

The processing of information in the stages of word production requires a certain amount of time. Once Indefrey and Levelt determined the cerebral localization of the component processes in their model, they proposed that the temporal properties of the regional activations should be compatible with timing estimates from previous electromagnetic studies (Levelt et al., 1998, Van Turennout, et al., 1997, 1998). Estimates from these studies implicate a time window between 275 and 400 msec for the lexical phonological code retrieval and phonological encoding processing components. Indefrey and Levelt suggested that the data available on the time course of word production is supported by the temporal sequence of processing components outlined in their model.

### 4.4. Summary

The word production model proposed by Indefrey and Levelt (2000) clearly distinguishes the various stages of processing involved in the generation of speech. In their model, processing begins with the combination of roots and derivational/inflectional information in the lexicon. From the lexicon, word forms are projected into the syntax to undergo syntactic operations and the subsequent insertion of morphophonological information. With such a detailed account of these processes, this model provides a framework that illustrates how cognitive tasks map directly onto linguistic representations. As a result, this model allows researchers to interpret perceptual processing studies in terms of a theoretical model of word production. Indefrey and Levelt applied this model to the data from a group of neuroimaging studies and observed significant activation patterns within a network of cortical regions corresponding to the processing components of their model. The present study extends this model to data from research involving fMRI and MEG neuroimaging methodology. The goal is to determine if the neural correlates and temporal sequencing of phonological processing identified by Indefrey and Levelt hold across tasks using alternative neuroimaging techniques.

## 5. fMRI and MEG Studies of Phonological Processing

The purpose of the present research is to determine if Indefrey and Levelt's model of word production can account for inconsistencies in the results of fMRI and MEG studies of phonological processing. To this end, I will integrate the results of fMRI and MEG research and compare them to the word production model to determine if the regions of activation from these studies match the cortical network for the processing components of this model.

### 5.1. fMRI Research

Phonological processing has received much attention in the functional neuroimaging literature. A number of studies using fMRI methodology show different patterns of activation for the computation of phonological information.

Shaywitz et al. (1995) reported activation restricted to the superior aspect of the inferior frontal gyrus and concluded that this area was uniquely associated with phonological processing. In a similar study, Pugh et al. (1996) observed activation in the inferior frontal gyrus and the middle and superior temporal gyri for the phonological task and concluded that phonological processing makes the heaviest demands on these regions of the cortex. Poldrack et al. (1999) determined that the posterior and dorsal region of the inferior frontal gyrus may be specialized for phonological processing. Although Lurito et al. (2000) observed some frontal activation, localized to Broca's area, their results also showed significant activation in temporal regions. Burton et al. (2000) found activation in superior temporal regions and determined that phonological processing does not necessarily recruit frontal areas. Based on the differences in the fMRI data, it is evident that there are still many potential locations for the specialization of phonological processing (see Table 2). The present research predicts that a reanalysis of this data, using Indefrey and Levelt's model of word production, will reveal that these studies have captured different aspects of phonological processing that correspond to the phonological components in this model.

(9) Table 2.

| Study | Task Type | Stimulus Type | Modality | Area of Activation |
|---|---|---|---|---|
| 1. Shaywitz et al. (1995) | Rhyme Judgement Task | Non-Word Pairs | Visual | •Inferior frontal gyrus |
| 2. Pugh et al. (1996) | Rhyme Judgement Task | Non-Word Pairs | Visual | • Inferior frontal gyrus<br>• Middle, superior temporal gyri |
| 3. Poldrack et al. (1999) | Syllable Counting Task | Words and Non-Words | Visual | • Inferior frontal gyrus (posterior and dorsal regions) |
| 4. Lurito et al. (2000) | Rhyme Determination Task | Words | Visual | • Broca's area<br>• Supramarginal gyrus<br>• Superior middle temporal gyrus |
| 5. Burton et al. (2000) | Onset Discrimination Tasks | CVC Word Pairs | Auditory | • Superior temporal gyrus (Exp's 1 & 2)<br>• Frontal regions (Exp 2 only) |

### 5.1.1 Shaywitz et al. (1995)

The purpose of the Shaywitz et al. (1995) study was to localize the neural regions associated with the component processing operations for reading. In order to engage phonological processing specifically, they used a rhyme judgement task with visually presented pseudoword stimuli. Participants were simultaneously presented with a pair of nonsense letter strings (e.g., LEAT and JETE) and were required to judge whether or not these strings rhymed using a button press response. Shaywitz et al. included an orthographic control condition to account for the influence of orthographic stimulus features on the phonological task. In this condition, the judgement decision was based on the case alternation pattern of a sequence of letters (e.g., <ttTt> and <tTtt>) in two visually presented consonant strings. Shaywitz et al. subtracted the orthographic control condition from the phonological condition using a subtraction paradigm and argued that the resulting pattern of activation reflected the neural substrates of phonological processing.

In this study, the experimental task requires the participants to respond to pseudoword stimuli. Because these stimuli would not correspond to any existing lexical entry, this task would not make any demands on semantic processing. To complete the task, participants would necessarily recruit knowledge about spelling-to-sound correspondences, in order to determine if the orthographic strings corresponded to rhyming syllables. According to the word production model, tasks that involve pseudoword reading engage the lead-in process of grapheme/phoneme conversion. This process accesses the core components of the word production system at the level of phonological encoding. In the word production model, phonological encoding is the second core phonological component, following phonological code retrieval (see Figure 6, pg. 22). The application of this model allows the phonological task in the Shaywitz et al. study to be mapped onto the processing component representing phonological encoding operations. In the cortical network for word production proposed by Indefrey and Levelt, phonological encoding should generate activation in the left inferior frontal gyrus and the left mid superior temporal gyrus.

Shaywitz et al. reported that, after subtraction, the pattern of activation for the phonological task was localized to the superior aspect of the inferior frontal gyrus, including Broca's area, and concluded that these regions are uniquely associated with phonological processing. The localization of phonological processes to the inferior frontal gyrus corresponds to the regions of activation for phonological encoding predicted by Indefrey and Levelt (see Table 3). However, Shaywitz et al. did not observe activation in the mid superior temporal gyrus. A possible explanation is that their task did not involve an auditory component. In the word production model, the core component of phonological encoding may be triggered by the lead-in processes of either auditory word repetition or grapheme/phoneme conversion. The cortical involvement of the mid superior temporal gyrus may be exclusive to the auditory lead-in process[1] and, since the phonological task in the Shaywitz et al. study recruited processes including grapheme/phoneme correspondences, activation may have been restricted to the inferior frontal gyrus.

(10) Table 3.

| Predicted Localizations | Area of Activation |
|---|---|
| ▪ Left inferior frontal gyrus | ✓ |
| ▪ Left mid superior temporal gyrus. | ✗ |

**5.1.2. Pugh et al. (1996)**

In a similar study, Pugh et al. (1996) used the same experimental paradigm as Shaywitz et al. (1995), with the added contribution of a visual-spatial control. They created a hierarchy of task conditions, each building incrementally on the previous condition. The first was a line-judgement task, designed to activate visual-spatial processing, where participants viewed two sets of four lines, one on top of the other, with right or left orientations and had to judge whether or not the lines displayed the same orientation pattern (e.g., \\\/ and \\/\). The second was a letter case judgement task with consonant string stimuli (e.g., <ttTt> and <tTtt). Pugh et al. argued that this condition makes demands on both visual-spatial and orthographic processing. In order to specifically activate phonological processes, Pugh et al. conducted a rhyme judgment task identical to the phonological task in the Shaywitz et al. study. Participants were presented with two nonsense word strings (e.g., REPE and MEAP) and were required to decide whether or not the nonword strings rhymed. This task was assumed to involve visual-spatial, orthographic and phonological processing. Pugh et al. also included a fourth, semantic, condition designed to involve all component operations, from visual-spatial processing up to semantic processing. Pugh et al. systematically subtracted each condition from the one above it in the hierarchy. To isolate the activation for phonological processing, they subtracted the case judgement task from the phonological task, with the assumption that this would eliminate interference from visual-spatial and orthographic processing.

Like the previous study, the experimental task in this study requires participants to access spelling-to-sound correspondences in order to generate phonological representations for the pseudowords. According to the word production model, this task would engage the lead-in process of

[1]This assumption may be possible to confirm from Indefrey and Levelt's statistical analysis in their study.

grapheme/phoneme conversion, which would activate the core process of phonological encoding. In this framework, the resulting pattern of activation should occur in the left posterior inferior frontal gyrus (e.g., Broca's area) and the left mid superior temporal gyrus.

Consistent with this prediction, Pugh et al. observed activation for the phonological task in frontal regions, including the inferior frontal gyrus and notably, some activation was reported for the middle and superior temporal gyri (see Table 4). Pugh et al. concluded that the neural correlates of phonological processing involve both frontal and temporal regions of the cortex. Activation in the inferior frontal gyrus confirms the localization of phonological encoding based on the word production model. Despite highly similar experimental designs, unlike the Shaywitz et al. study, Pugh et al. reported activation in the medial and superior temporal gyri. These results are somewhat surprising considering that Shaywitz et al. used essentially the same phonological task and reported no temporal involvement. The possibility that the temporal localizations described in Indefrey and Levelt's cortical network correspond only to the auditory lead-in process for phonological encoding does not seem to hold, as the phonological task in this study did not involve the auditory modality.

(11) Table 4.

| Predicted Localizations | Area of Activation |
|---|---|
| ▪ Left inferior frontal gyrus | ✓ |
| ▪ Left mid superior temporal gyrus. | ✓ |

**5.1.3. Poldrack, et al. (1999)**

Poldrack et al. (1999) conducted a study designed to examine the role of the left inferior prefrontal cortex in semantic and phonological proccssing, using two different experimental phonological conditions. Both of these conditions were subtracted from a baseline perceptual task to isolate phonological processing. The phonological task in the first condition involved counting the syllables for visually presented mono-, di- or tri-syllabic words and the orthographic task was a case judgement task that required a decision response about the case of the presented words (e.g., uppercase vs. lowercase). The second scan condition compared another phonological syllable counting task, this time with pseudoword stimuli, to the same case judgement orthographic task.

In this study, the experimental task for the first phonological condition recruits a number of phonological operations. Prior to the computation of syllable structure information, participants presumably identify the stimulus as a word form. According to the word production model, word reading tasks engage the lead-in process of visual word recognition. Inherent to this component is the assumption that word recognition is automatic, even when not explicitly required to complete the task. Support for this assumption comes from other types of cognitive tasks (e.g., the Stroop Task). This process mediates the word production system at the level of phonological retrieval. This step is required because, according to Indefrey and Levelt, word recognition would trigger the retrieval of the lemma associated with each lexical concept, as these lemmas are the characteristic input of the phonological code retrieval processing component. Once selected, these lemmas generate the phonological codes of the word and it is these codes that provide the input for the syllabification of the word. The syllable is one of the constituent units in the core processing component of phonological encoding, so this component would also be activated by the syllable counting task. Consequently, the phonological task in the first phonological condition would involve all levels of phonological processing specified in the word production model. As a result, cortical activation should be widespread in this condition of this experiment. Based on the predictions from Indefrey and Levelt, this activation should involve the left posterior superior and middle temporal gyri (e.g., Wernicke's area) and the left thalamus for phonological retrieval and the left posterior inferior frontal gyrus (e.g., Broca's area) and the left mid superior temporal gyrus for phonological encoding.

The second phonological condition of the experiment also required syllable counting, but this condition used pseudoword stimuli. Unlike the first condition, this condition would not generate the retrieval of phonological code information, as these codes are only associated with items stored in the lexicon. As a result, the phonological codes that provide the input to syllabification would be generated strictly from the translation of orthographic features into phonological features. According to the word production model, pseudoword reading tasks engage the lead-in process of grapheme/phoneme conversion, which, in turn, triggers phonological encoding. As a result, the pattern of activation for the second condition should be restricted to the left posterior inferior frontal gyrus (e.g., Broca's area) and the left mid superior temporal gyrus.

Unfortunately, since the focus of this study was on the role of the inferior prefrontal cortex in language processing, Poldrack et al. restricted their analysis to the scanning range of frontal regions. Despite this fact, the pattern of results observed in this study still appear to be consistent with the localizations found by Indefrey and Levelt in their meta-analysis. For both phonological task conditions, Poldrack et al. observed activation in the dorsal portions of the left inferior frontal gyrus and for the second task, activation was also significant in the posterior sections of the inferior frontal gyrus (e.g., Broca's Area) (see Table 5). These results are consistent with the frontal region correlates of phonological encoding outlined in the cortical network for word production processes. Poldrack et al. conducted no analysis for activation outside of these frontal regions and, as a result, is it impossible to determine if the patterns of activation produced by these tasks correspond to the localizations predicted for the other components of the word production model.

(12) Table 5.

| Predicted Localizations | Area of Activation | |
|---|---|---|
| | Word Task | Pseudoword Task |
| ▪ Left inferior frontal gyrus | ✓ | ✓ |
| ▪ Left mid superior temporal gyrus | N/A | N/A |
| ▪ Left posterior, superior and middle temporal gyri | N/A | N/A |
| ▪ Left thalamus | N/A | N/A |

Importantly, however, Poldrack et al. noted that there was much less activation of frontal areas for the syllable counting task involving real word stimuli. They suggested that this was the likely result of activation in posterior regions (e.g., regions in temporal cortex) outside of the scanning range set for this experiment. Poldrack et al. suggested that these regions may have played a primary role in task performance, minimizing the involvement of frontal cortex. This explanation is possible in the context of the word production model, as word reading tasks activate phonological retrieval operations, which are associated with activation in the left posterior superior and middle temporal gyri. However, without activation data for regions outside of the frontal cortex, this explanation is difficult to confirm.

**5.1.4. Lurito et al. (2000)**

In an analysis of the processes involved in word generation, Lurito et al. (2000) used a silent word rhyming paradigm to activate phonological processing. This paradigm consisted of a rhyme judgment task that was designed to generate subvocal articulation by using orthographically different rhyming word pairs (e.g., 'dial' and 'file') and orthographically similar non-rhyming word pairs (e.g., 'comb' and 'tomb') as stimuli. Participants were required to determine if the two visually presented common English words rhymed and then make a button press response. The control task was a line judgement task, where participants had to judge if a set of lines in different orientations were identical or not (e.g., [\\/\] and [/\\\]). Lurito et al. summed the single subject activation maps and activations were considered to be significant only if they exceeded a set threshold in the combined map.

The phonological task in this study was a task of basic word reading. The use of common English words for the rhyme judgment recruits word recognition processes. According to the word production model, word reading invokes the lead-in process of visual word recognition. This process enters the word production system at the core processing level of phonological code retrieval. The pattern of cortical activation associated with phonological retrieval is in the left posterior superior and middle temporal gyri (e.g., Wernicke's area) and the left thalamus.

Lurito et al. reported significant activation in the supramarginal gyrus (superior temporal cortex) and a band of activation deep within the superior region of the left middle temporal gyrus. They concluded that the rhyming task made the greatest demands on peri-sylvian language areas (e.g., Wernicke's area). Activation in these temporal regions confirms the localization of phonological retrieval operations based on the word production model (see Table 6). Although Indefrey and Levelt predicted the activation in the left thalamus, most analyses do not record subcortical structure activation.

(13) Table 6.

| Predicted Localizations | Area of Activation |
|---|---|
| ▪ Left posterior, superior and middle temporal gyri | ✓ |
| ▪ Left thalamus | N/A |

**5.1.5. Burton et al. (2000)**

To characterize the nature of activation in segmental processing, Burton et al. (2000) conducted two experiments involving 'same/different' judgements for auditorily presented word pair stimuli. In the first experiment, the 'different' word pairs only differed in the voicing feature of the onset (e.g., [dip] and [tip]). As a result, Burton et al. argued that these pairs do not require overt segmentation. By contrast, the second experiment had word pairs that differed by onset voicing and vowel and coda phonemes (e.g., [dip] and [ten]). Burton et al. argued that these pairs required participants to segment the stimuli. Both experiments were subtracted from a tone-discrimination control condition to reveal the regions of activation.

These experiments both involve the computation of phonological information from an auditory speech source. Burton et al. argued that the first experiment requires only the perception of the words to complete the judgement task, but that the second experiment involves the additional step of segmenting the speech stream into individual phonemes. In the word production model, there are two lead-in processes associated with phonological encoding: auditory word perception and grapheme/phoneme conversion. In this study, the first experiment only involves the lead-in process of auditory word perception. The second experiment requires the recruitment of processes in addition to auditory word perception, but the processes related to speech segmentation are also associated with phonological encoding operations. In terms of the cortical network proposed in the Indefrey and Levelt study, the results from this study should show activation in the left posterior inferior frontal gyrus (e.g., Broca's area) and the left mid superior temporal gyrus.

Burton et al. reported differential patterns of activation for these two experiments. Both speech conditions displayed activation in superior temporal regions, but only the second condition showed additional activation in frontal areas. From these findings, Burton et al. concluded that phonological processing does not necessarily always recruit frontal areas. This pattern of results is corresponds to the regions specified for phonological encoding in the word production model (see Table 7). The first task only involved auditory word perception and, as such, activation was limited to temporal regions. The second experiment involved additional processing operations which resulted in activation for all areas associated with phonological encoding.

(14) Table 7.

| Predicted Localizations | Area of Activation |
|---|---|
| ▪ Left inferior frontal gyrus | ✓ |
| ▪ Left mid superior temporal gyrus | ✓ |

The application of Indefrey and Levelt's word production model to the data from this sample of fMRI studies shows that, using a theoretical model of the processes involved in lexical retrieval, it is possible to reliably identify the neural correlates of phonological processing. Although some of the results were not consistent with this model, the majority of findings converged onto the cortical network proposed by Indefrey and Levelt. In the next section, I will apply this model to a group of MEG studies in an attempt to obtain a similar result.

**5.2. MEG Research**

MEG has yet to become widely applied to the investigation of language processing. However, in the existing literature using MEG methodology to test the phonological characteristics of language function, there are subtle differences in the patterns of results for phonological processing. Zouridakis et al. (1998) determined that the source of activity was localized to temporal and temporoparietal areas between 200 and 600 msec post-stimulus onset. Simos et al. (1998) observed that the activity sources in their study were localized to the superior and middle temporal gyri, however, differential patterns of response latencies were observed across modalities. Phillips et al. (2000) reported that the localization of their mismatch response was centered in the superior temporal cortex, in the 150-210 msec time window. The results of Phillips et al. (in press) were localized to the supratemporal gyrus in the left hemisphere with the divergence of responses occurring at a series of latency intervals. Evidence from MEG research shows the overall convergence of activation patterns for phonological processing on regions in the temporal cortex, however, differences arise in the temporal properties of this activation (see Table 8.0). The application of the word production model to this data may confirm the localization of phonological processes within the temporal cortices and clarify results associated with the temporal sequencing of such processes.

(15) Table 8.0

| Study | Task Type | Stimulus Type | Modality | Area/Timing of Activation |
|---|---|---|---|---|
| 1. Zouridakis et al. (1998) | Word Identification Task | Word Lists | Visual | • Temporal and temporoparietal regions.<br>• 200-600 msec |
| 2. Simos et al. (1998) | 1. Auditory Word Recognition Task<br><br>2. Visual Word Recognition Task | Words | Auditory/ Visual | • Superior and middle temporal gyri<br>• Auditory: 100, 210, 350 msec<br>• Visual: 350-500 msec |
| 3. Phillips et al. (2000) | Passive Listening Task | Phoneme Sequences | Auditory | • Superior temporal cortex<br>• 150-210 msec |
| 4. Phillips et al. (in press) | Identification Task | Phoneme Sequences | Auditory | • Supratemporal gyrus<br>• 170-230 msec |

**5.2.1. Zouridakis et al. (1998)**

In order to investigate the degree of hemispheric activation during tasks of language function, Zouridakis et al. (1998) conducted a word identification task using real word stimuli. Participants were required to read a list of real words and identify familiar words in a subsequent recognition memory task. The control condition involved a face recognition task, where participants viewed a series of unfamiliar human face photographs and then completed a recognition memory test. For each condition, the observed activity was indexed by single equivalent current dipoles (ECD) and a dipole localization algorithm determined which activations were considered significant.

In this study, the word identification task was a task of basic word reading. This task would involve the retrieval of information associated with each lexical form. According to the word production model, word reading engages the lead-in process of visual word recognition. This lead-in process activates the core processing system at the level of phonological code retrieval. This process would involve the selection of a lemma for each lexical concept in the list of word stimuli. Information in the selected lemmas would be used to generate the phonological code for each word. It is this code that would subsequently be recalled for the recognition memory task. Based on the cortical network proposed by Indefrey and Levelt, the pattern of activation for phonological code retrieval should involve the left posterior superior and middle temporal gyri (e.g., Wernicke's area) and the left thalamus and the timing estimates confirmed by the localizations in the Indefrey and Levelt study predict a time window between 275 and 400msec for phonological code retrieval.

Zouridakis et al. identified the localization of dipole activation to superior and middle temporal gyri and temporo-parietal regions for the word recognition task. The average MEG responses for these activations showed that ECDs associated with activation in language areas occurred between 200 and 600 msec after the stimulus onset. The activation in the temporal gyri is consistent with the regions of activation for phonological code retrieval described in the word production model (see Table 9). Although the timing sequence falls mainly within the window predicted by the timing estimates in the Indefrey and Levelt study, responses at the later intervals (>400msec) exceed the bounds of these predictions.

(16) Table 9.

| Predicted Localizations and Time Course | Area of Activation |
|---|---|
| ▪ Left posterior, superior and middle temporal gyri | ✓ |
| ▪ Left thalamus | ✖ |
| ▪ 275 - 400msec | ✖ |

**5.2.2. Simos et al. (1998)**

Simos et al. (1998) conducted an experiment to explore which cortical areas underlie language comprehension processes. Participants were tested on two continuous recognition memory tasks, the first involving visual word stimuli and the second in the auditory modality. Participants either viewed or listened to a word stimuli list and were required to identify words from the test list in a subsequent recognition memory task. Similar to the previous study, the control condition involved a face recognition task with photographs of unfamiliar human faces used as stimuli.

In this study, the visual task condition triggers the lead-in process of visual word recognition. This process would recruit phonological code retrieval operations, which according to the cortical network for word production, would result in patterns of activation in the left posterior superior and middle temporal gyri (e.g., Wernicke's area) and the left thalamus. The auditory task condition activates the lead-in process of auditory word perception. This process generates phonological encoding processes and, as such, would produce activation in the left posterior inferior frontal gyrus (e.g., Broca's area) and the left mid superior temporal gyrus. Again, the timing estimates from the Indefrey and Levelt study predict responses at 275 and 400msec for both phonological code retrieval and phonological encoding component processes.

For the visual recognition task, Simos et al. reported activation in the superior and middle temporal gyrus and at the temporo-parietal junction. These responses developed between 350 and 500 msec after the onset of the stimulus. For the auditory recognition task, Simos et al. reported activation in the vicinity of the superior and middle temporal gyri. Notably, activation in prefrontal regions was observed for two participants. The auditory responses occurred at three early latency intervals corresponding to 100, 210 and 350 msec post-stimulus onset. The results for both conditions in this experiment converge on the localization evidence based on the word production model. The latency interval for the visual task slightly exceeds the timing estimates proposed by Indefrey and Levelt and the responses in the auditory condition are earlier than predicted time window.

(17) Table 10.

| Predicted Localizations and Time Course | Area of Activation |
|---|---|
| ▪ Left inferior frontal gyrus | **? (2 participants)** |
| ▪ Left mid superior temporal gyrus | ✓ |
| ▪ Left posterior, superior and middle temporal gyri | ✓ |
| ▪ Left thalamus | ✖ |
| ▪ 275 - 400msec | ✖ |

**5.2.3. Phillips et al. (2000)**

Phillips et al. (2000) used a modified mismatch paradigm to determine if phonological category representations are available in auditory cortices. There were two experiments in this study, the first was a phonological condition to determine if the auditory cortex can access phonological category representations and the second was an acoustic condition that manipulated the perceptual category boundary of the stimuli from the main (e.g., phonological) condition. Participants listened passively to a continuum of speech sounds. The phonological condition contrasted stimuli from a /dæ/ - /tæ/ continuum and, in the acoustic condition, the phonological distribution of the stimuli was changed, such that the perceptual category boundary fell between the second and third stimulus.

Both conditions of this study involved a passive listening task for the discrimination of phonological category boundaries. This experiment does not involve the same types of processes as in the previous MEG studies. However, due to the lack of literature using MEG techniques to determine cortical activation for phonological information, I have included these studies, comparing them to processes present in the word production model. In the word production model, there is no lead-in process that corresponds specifically to passive listening tasks, however, the auditory perception of a continuum of phonological stimuli would presumably be comparable to auditory word perception. The lead-in process of auditory word perception activates the core processing component of phonological encoding. In terms of the cortical network proposed in the Indefrey and Levelt study, the results from this study should show activation in the left posterior inferior frontal gyrus (e.g., Broca's area) and the left mid superior temporal gyrus between 275 and 400msec for phonological encoding operations.

Phillips et al. (2000) reported an early mismatch negativity response in the 150-210 msec latency window localized to auditory cortex. They suggested that their findings support the notion that the phonetic category of voicing is localized to the auditory cortex. Although Phillips et al. did not report activation in frontal regions, their findings do correspond to the temporal regions of the cortical network for word production (see Table 11). The time course of their mismatch response is markedly earlier than the timing estimates provided by the Indefrey and Levelt study.

(18) Table 11.

| Predicted Localizations and Time Course | Area of Activation |
|---|---|
| ▪ Left inferior frontal gyrus | ✖ |
| ▪ Left mid superior temporal gyrus | ✓ |
| ▪ 275 - 400msec | ✖ |

**5.2.4. Phillips et al. (in press)**

In a similar study, Phillips et al. (in press) also used an auditory mismatch paradigm to localize the neural regions associated with phonological classes. Participants initially completed an identification task to establish the perceptual boundaries for each subject across three place of articulation categories (e.g., labial, alveolar, and velar). Again a passive listening task was employed to determine if this category distinction was supported by the auditory cortex. In the phonological feature condition, participants listened to stimuli from these categories randomly dispersed with deviant stimuli. A second acoustic control condition was run to determine if a mismatch response for the first condition could be explained by acoustic properties of the stimuli, rather than properties of the phonological feature categories.

Again, these tasks would activate the lead-in process of auditory word perception. This process would recruit phonological encoding operations. According to the word production model, activation for this operation should be localized to the left posterior inferior frontal gyrus (e.g., Broca's area) and the left mid superior temporal gyrus. The timing estimates from the Indefrey and Levelt study predict that a mismatch response for phonological encoding processes would occur in a latency window between 275 and 400 msec.

Similar to the previous study, Phillips et al. observed a mismatch response at a 170-230 ms time interval. The divergence of responses occurred across the anterior and posterior channel groups, indicating that the generator of the response was localized in the supratemporal auditory cortex. Again, Phillips et al. did not report activation in frontal regions, however, the patterns of results in temporal cortex remains consistent with cortical network from Indefrey and Levelt's study (see Table 12). The latency interval of the mismatch response in this study is approaching the Indefrey and Levelt's time estimates, but are still slightly earlier than the predicted latency response.

(19) Table 12.

| Predicted Localizations and Time Course | Area of Activation |
|---|---|
| ▪ Left inferior frontal gyrus | ✖ |
| ▪ Left mid superior temporal gyrus | ✓ |
| ▪ 275 - 400msec | ✖ |

Similar to the results from the fMRI research, the MEG data for the localization of phonological processing was also consistent with the word production model proposed by Indefrey and Levelt. Although the majority of MEG responses were centered in regions that corresponded to the cortical network for word production, most of the latency results for these responses did not support the timing estimates from the Indefrey and Levelt study. It is likely that the relatively small number of studies in this analysis were not representative of overall latency measures for phonological processing in the MEG literature and it is possible that the analysis of a broader sample of results may show a convergence of latency data with the estimates from Indefrey and Levelt's study.

## 6. Conclusions, Implications and Directions for Future Research

The investigation of phonological processes using neuroimaging techniques has produced variable results. Poeppel (1996) proposed that much of the variability in PET data investigating phonological processes could be accounted for by reanalyzing the results in terms of a theoretical model of lexical representation. Indefrey and Levelt (2000) applied their model of word production to the results of a sample of mainly PET neuroimaging studies and observed the convergence of these results onto a specific network of cortical regions. They concluded that their model of word production provides a theoretical framework for the analysis of neuroimaging research on the processes involved in the generation of words.

The purpose of the present study was to apply this model of word production to the results of a group of fMRI and MEG studies to test if these results also corresponded to the cortical network for word production proposed by Indefrey and Levelt (2000). The application of the word production model to results from fMRI and MEG research revealed that the regions of activation in these studies were consistent with the regions outlined in the cortical network from the Indefrey and Levelt study. These results demonstrate that, with this model, it is possible to reliably determine the neural correlates of phonological processing.

The implications of these results for Indefrey and Levelt's theory of word production model are that their processing model can be used as a guiding theoretical framework for the study of phonological processing using neuroimaging techniques. The convergence of evidence from the fMRI and MEG data provide support for the sequence of processing components outlined in this model. In their meta-analysis, Indefrey and Levelt determined that their model provided a theoretical structure for the resolution of inconsistent PET and ERP data on language processing. This study has extended the application of this model to fMRI and MEG research investigating specifically processes associated with the phonological components of the model and confirmed that the resolution of divergent fMRI and MEG data is also possible. Although the results of the fMRI and MEG studies in this analysis were predominantly consistent with the model of word production proposed by Indefrey and Levelt, some of the data remains problematic.

Some of the data from these studies did not converge as expected onto the cortical network proposed in the Indefrey and Levelt study. Although the cortical network for word production predicts activation in both the inferior frontal gyrus and mid superior temporal gyrus for phonological encoding operations, it was rare to find activation in both of these areas for these types of tasks. In fact, only two studies from this analysis reported activation in both of these regions when the lead-in processes of auditory word perception and grapheme-phoneme conversion were activated (e.g., Pugh, et al., 1996; Burton, et al., 2000). Similarly, although Indefrey and Levelt propose activation in the left thalamus for phonological code retrieval processes, none of the studies in this analysis reported subcortical activation in the thalamic regions. The timing estimates from the Indefrey and Levelt study also proved problematic for the MEG research in this study. Indefrey and Levelt predicted latency windows of between 275 and 400 msec for both phonological code retrieval and phonological encoding processing components. Although the latency response data from the MEG studies all showed activation in time windows close to these estimates, most responses were either slightly earlier or later than the window predicted in the Indefrey and Levelt study. It appears that, although the model of word production proposed by

Indefrey and Levelt accounts for a significant amount of the variability in neuroimaging results, some problems still remain. One possibility for the resolution of remaining issues is the consideration of alternative accounts of language processing.

Psycholinguistic theories of the speech processing propose an interactive system for the generation of complex word forms. Dell (1988) proposed a model of phonological retrieval in which the connections between word and phoneme nodes are bi-directional, allowing feedback between the various levels of representation in the network. This spreading activation allows activated word nodes to retrieve their corresponding constituent phonemes in the lexical network. Connectionist models of phonological processing present an alternative to strictly modular accounts of lexical representation for future research on the cortical basis of phonological processing. Perhaps with greater knowledge of the interactivity between the various types of processes in word production, there can be a better understanding of the neural substrates of phonological processing.

**Acknowledgements**
Special thanks to my supervisor Martha McGinnis for all her patience and guidance with this project and to my second reader, Craig Chambers, for all of his advice. Thanks also to Alec Marantz and David Poeppel for their time and helpful comments.

**References**

Anderson, S.R. (1992). *A-Morphous Morphology.* Cambridge, MA, USA: Cambridge University Press.

Baddeley, A. (1992). Working memory. *Science, 255,* 556-559.

Burton, M.W., S.L., Small, S.E. Blumstein (2000). The role of segmentation in phonological processes: An fMRI investigation. *Journal of Cognitive Neuroscience, 12(4),* 679-690.

Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale & S.J. Keyser, (Eds.) *The View from Building 20.* Cambridge, MA, USA: The MIT Press. pp. 1-52.

Demb, J. B., R.A. Poldrack & J.D.E. Gabrieli (1999). Functional neuroimaging of word processing in normal and dyslexic readers. In K. M. Raymond, P. A. McMullen (Eds.) *Converging methods for understanding reading and dyslexia. Language, speech, and communication* (pp. 245-304). Cambridge, MA, USA: The MIT Press.

Demonet, J.F., F. Chollet, S. Ramsay, D. Cardebat, J.L. Nespoulous, R. Wise, A. Rascol, & R. Frackowiak, (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain, 115,* 1753-1768.

Demonet, J.F., R. Wise, & R. Frackowiak, (1993). Language functions explored in normal subjects by positron emission tomography: A critical review. *Human Brain Mapping, 1,* 39-47.

Fox, P.T., M.A. Mintun, E.M. Reiman, & M.E. Raichle, (1988). Enhanced detection of focal brain responses using inter-subject averaging and change-distribution analysis of subtracted PET images. *Journal of Cerebral Blood Flow and Metabolism, 8,* 642-653.

Friston, K., C. Frith, & R. Frackowiak, (1993). Time-dependent changes in effective connectivity measured with PET. *Human Brain Mapping, 1,* 69-79.

Halle, M., & A. Marantz, (1993). Distributed morphology and the pieces of inflection. In Hale, K., & S.J. Keyser, (Eds.) *The View from Building 20.* Cambridge, MA, USA: The MIT Press, 111-176.

Hämäläinen, M., R. Hari, R.J. Ilmoniemi, J. Knuutila, & O.V. Lounasmaa, (1993). Magnetoencephalography – theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics, 65,* 413-497.

Indefrey, P., & W.J.M. Levelt, (2000). The neural correlates of language production. In M. Gazzaniga (Ed.) *The Cognitive Neurosciences, Second Edition,* 845-865. Cambridge, MA, USA: The MIT Press.

Jaeger, J.J., A.H. Lockwood, D.L. Kemmerer, R.D. Van Valin, B.W. Murphy, & H.G. Khalak, (1996). A positron emission tomographic study of regular and irregular verb morphology in English. *Language, 72 (3),* 451-497.

Klein, D., B. Milner, R.J. Zatorre, E. Meyer, & A.C. Evans, (1995). The neural substrates underlying word generation: A bilingual functional-imaging study. *Proceedings of the National Academy of Sciences USA, 92,* 2899-2903.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation.* Cambridge, MA, USA: The MIT Press.

Levelt, W.J.M., P. Praamstra, A.S. Meyer, P., Helenius, & R., Salmelin (1998). An MEG study of picture naming. *Journal of Cognitive Neuroscience, 10,* 553-567.

Levelt, W. J. M., & P. Indefrey, (2000). The speaking mind/brain: Where do spoken words come from? In Marantz, A., Y. Miyashita, W. O'Neil (Eds.) *Image, Language, Brain. Papers from the First Mind Articulation Project Symposium.*

Lieber, R. (1992). *Deconstructing Morphology. Word Formation in Syntactic Theory.* Chicago, IL: The University of Chicago Press.

Lurito, J.T., D.A. Kareken, M.J. Lowe, S.H.A Chen, V.P. Mathews, (2000). Comparison of rhyming and word generation with fMRI. *Human Brain Mapping, 10,* 99-106.

Ogawa, S., T.M. Lee, A.S. Nayak, & P. Glynn, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *MagneticResonance in Medicine, 14,* 68-78.
Ogawa, S., D.W. Tank, R. Menon, J.M. Ellermann, S.G. Kim., H. Merkle, & K. Ugurbil, (1992). Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences USA, 89,* 5951-5955.
Paulesu, E., C.D. Frith, & R.S.J. Frackowiak, (1993). The neural correlates of the verbal component of working memory. *Nature, 362,* 342-345.
Petersen, S., P. Fox, M. Posner, M. Mintun, & M. Raichle, (1989). Positron emission tomographic studies of the processing of single words. *Journal of Cognitive Neuroscience, 1(2),* 153-170.
Phillips, C., A. Marantz, M. McGinnis, D. Petesky, K. Wexler, E. Yellin, D. Poeppel, T. Roberts, & H. Rowley, (1995). Brain mechanisms of speech perception: A preliminary report. In C., Schütze, J., Ganger, & K. Broihier, (Eds.) *Papers on Language Processing and Acquisition. MIT Working Papers in Linguistics, #26.* Department of Linguistics and Philosophy, MIT.
Phillips, C., T., Pellathy, A., Marantz, E., Yellin, K., Wexler, D., Poeppel, M., McGinnis, & T., Roberts, (2000). Auditory cortex accesses phonological categories: An MEG mismatch study. *Journal of Cognitive Neuroscience, 12(6),* 1-18.
Phillips, C., T., Pellathy, A., Marantz, (submitted, July 2000). Phonological feature representations in auditory cortex.
Poeppel, D. (1996). A critical review of PET studies of phonological processing. *Brain and Language, 55,* 317-351.
Poldrack, R.A., A.D., Wagner, M.W., Prull, J.E., Desmond, G.H., Glover, & J.D.E., Gabrieli, (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *NeuroImage, 10,* 15-35.
Posner, M., S., Petersen, P., Fox, & M., Raichle, (1988). Localization of cognitive operations in the human brain. *Science, 240,* 1627-1631.
Price, C.J., C.J., Moore, G.W., Humphreys, & R.J.S., Wise, (1997). Segregating semantic from phonological processes during reading. *Journal of Cognitive Neuroscience, 9,* 727-733.
Pugh, K.R., B.A., Shaywitz, S.E., Shaywitz, T., Constable, P., Skudlarski, R.K., Fulbright, R.A., Bronen, D.P., Shankweiler, L., Katz, J.M., Fletcher, & J.C., Gore, (1996). Cerebral organization of components processes in reading. *Brain, 119,* 1221-1238.

Rugg, M.D. (1997). Functional neuroimaging in cognitive neuroscience. In M.D. Rugg (Ed.) *Cognitive Neuroscience,* (pp. 15-36). Cambridge, MA, USA: The MIT Press.

Sergent, J., E., Zuck, M., Levesque, & B., MacDonald, (1992). Positron emission tomography study of letter and object processing: Empirical findings and methodological considerations. *Cerebral Cortex, 2,* 68-80.

Shaywitz, B.A., S.E., Shaywitz, K.R., Pugh, T., Constable, P., Skudlarski, R.K., Fulbright, R.A., Bronen, J.M., Fletcher, D.P., Shankweiler, L., Katz, & J.C., Gore, (1995). Sex differences in the functional organization for language. *Nature, 373,* 607-609.

Simos, P.G., J.I., Breier, G., Zouridakis, & A.C., Papanicolaou, (1998). Identification of language-specific brain activity using magnetoencephalography. *Journal of Clinical and Experimental Neuropsychology, 20 (5),* 706-722.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643.661.

Van Turennout, M., P. Hagoort, & C.M., Brown (1997). Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23,* 787-806.

Van Turennout, M., P. Hagoort, & C.M., Brown (1998). Brain activity during speaking: From syntax to phonology in 40 milliseconds. *Science, 280,* 572-574.

Wunderlich, D. (1997). A Minimalist model of inflectional morphology. In C., Wilder, (Ed.) *The Role of Economy Principles in Linguistic Theory.* Berlin, Germany: Akademie, 267-298.

Zatorre, R., A., Evans, E., Meyer, & A., Gjedde, (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science, 256,* 846-849.

Zouridakis, G., P.G., Simos, J.I., Breier, & A.C., Papanicolaou, (1998). Functional hemispheric asymmetry assessment in a visual language task using MEG. *Brain Topography, 11,* 57-65.

Jodi Edwards
Department of Linguistics
University of Calgary
2500 University Drive NW
Calgary, AB T2N 1N4
Fax: (403) 282-3880
jdedward@ucalgary.ca

**The Acoustic Correlates of Blackfoot Prominence**

Sheena Van Der Mark
University of calgary

**Abstract**

Blackfoot, an Algonquian language spoken in Alberta and Montana, has been described as a pitch accent language (Frantz and Russell 1989; Frantz 1991; Kaneko 1999). Pitch accent languages mark phonetic prominence with a difference in pitch on the prominent syllable. Beckman (1986) has shown that Japanese (a prototypical pitch accent language) differs from English (a prototypical stress language) in that fundamental frequency (pitch) is the *only* variable that marks prominence in Japanese, whereas several variables mark prominence in English. These variables include fundamental frequency ($F_0$) peak, amplitude peak, average amplitude, total amplitude and duration. Based on Beckman's analysis of Japanese, we would expect Blackfoot, as a pitch accent language, to mark prominence only with $F_0$, thus patterning with Japanese. However, this analysis shows that in addition to $F_0$, average amplitude was also correlated with prominence in Blackfoot, amplitude peak, total amplitude and duration were not. These results suggest that Blackfoot is different than Japanese in how prominence is marked. However, the results are similar enough to justify the classification of Japanese as a pitch accent language.

## 1. Introduction

### 1.1. Objectives

The purpose of this thesis is to examine the acoustic correlates of prominence in Blackfoot. Blackfoot is claimed to be a pitch accent language (Frantz and Russell 1989; Frantz 1991; Kaneko 1999), that is, a language which indicates that phonetic prominence is marked by a difference exclusively or largely in pitch on the prominent syllable. However, to my knowledge, no acoustic analysis has ever been done to determine if this is in fact how Blackfoot marks prominence. The only in-depth examination of Blackfoot pitch accent was a phonological analysis (Kaneko) apparently based on subjective interpretations of pitch changes in words spoken by Native Blackfoot speakers from the Blood reserve, which is located between Lethbridge and Cardston. She conducted an Optimality Theory analysis of accent patterns in Blackfoot nominals. Although this work greatly contributed to our knowledge of the patterns of pitch accent in Blackfoot, the conclusions were purely phonological. The present study will look at Blackfoot prominence from a phonetic point of view, focusing on how it is manifested acoustically.

### 1.2. Outline of the Thesis

The rest of this chapter goes over the context and methodology of data collection and information on the consultants. Chapter two reviews the phonological and phonetic attributes of different accent systems. Section 2.1 discusses the definitions of terms used throughout this thesis. Phonological prominence is discussed in §2.3, and previous research on the acoustic correlates of phonological prominence (stress and pitch accent) in other languages is discussed in §2.4. Chapter three discusses the Blackfoot language, going over the phonemic inventory in §3.1, Blackfoot syllable structure in §3.2, and the pitch accent patterns as analyzed by Kaneko in §3.3. The acoustic analysis is presented in Chapter four, with the procedures discussed in §4.1. Five variables will be analyzed: fundamental frequency, duration, amplitude peak, average amplitude and total amplitude. The results will be presented in §4.2. Chapter five presents a statistical analysis of the data, determining the significance of the results found in chapter four. Statistical procedures and their interpretations are discussed in §5.1, and the results of the statistical analysis are presented in §5.2. Chapter six discusses the results and conclusions of the acoustic and statistical analyses. §6.2 summarizes the results of this paper and §6.3 discusses the next steps that need to be taken for further research on Blackfoot prominence.

### 1.3. The Blackfoot Language

Blackfoot is a Native American language belonging to the Algonquian family. It is thought to be the most divergent of all Algonquian languages, suggesting that it was the first to fully split from Proto-Algonquian, (Proulx 1989). Blackfoot is spoken in Alberta and Northern Montana by the Blood, Peigan and Siksika tribes. The Siksika reserve is about one-hundred kilometers southeast of Calgary, the Blood reserve is located between Cardston and Lethbridge, the Peigan reserve is located west of Fort MacLeod, and the Blackfeet reserve is in Northern Montana. A slightly different dialect is spoken on each of these reserves. Today there are approximately 5000 Blackfoot speakers.

The language is currently undergoing a lot of change, resulting in the separate designations Old Blackfoot and Modern Blackfoot. It is likely that there is no distinct break between these two varieties, but that they represent two abstract points on a continuum of change within the language.

Most, if not all, remaining Blackfoot speakers are bilingual speakers of English. The influence of English may well be one of the reasons for the rapid change occurring within the language.

### 1.3. Data Collection, Methodology and Consultant Information

The data analyzed in this thesis was elicited from native speakers of Blackfoot. Most of the elicitations were done from January to April 2000 in a Field Methods class at the University of Calgary, although a small number of words were elicited in April 2001 specifically for this study. The first consultant was PB, a female who grew up on the Siksika reserve speaking Blackfoot as her first language. The second consultant was PB's older sister NB.

PB and NB made a distinction between Old and Modern Blackfoot, which is also discussed in Kaneko. The consultants often distinguished between different lexical and morphological forms used for Old and Modern Blackfoot. In citing an Old Blackfoot term, they would often suggest that this is the term an elder would use. The consultants indicated that NB spoke Old Blackfoot, while PB spoke Modern Blackfoot. However, although some of the forms elicited were different, they were similar enough to categorize them both as Modern Blackfoot. This may simply reflect the fact that each of the consultants fall along different places on a continuum from one variety to the other.

All of the data elicited in class was recorded on a Sony Professional Walkman. Transcriptions made at the time of elicitations were discussed by members of the Field Methods class and later checked with the recordings of the sessions.

## 2. Overview of Accent Systems

### 2.1. Pitch, Tone and Stress: Definitions

In the linguistics literature, the terms tone language, pitch accent, and stress accent are often used to classify languages. However, it is not always made clear what is meant by each of these terms, and different authors often seem to mean different things. In order to clarify the meanings these terms are intended to represent, I will outline the definitions I will follow throughout this thesis.

First I note that *pitch* is the perceptual correlate of the acoustic variable of fundamental frequency ($F_0$).

*Tone languages* are languages in which each word or syllable carries lexically contrastive pitch. Differences in pitch correlate with differences in meaning even when all other phonemic information is identical.

Stress will be referred to as either phonological stress or phonetic stress. *Phonological stress* is a formal method of marking one syllable in a word as more prominent than others.[1] This is usually done through modulating pitch, which is how *pitch accent languages* manifest phonological prominence.

[1] In this paper, *phonological stress* will be used interchangeably with *prominence* and *accent*.

In *stress languages,* a phonologically prominent syllable is marked through modulating phonetic stress. Phonetic stress as a method of marking prominent syllables is not as easy to define as pitch accent, as it may involve several acoustic variables. The following sections will discuss these definitions in more detail, and the acoustic variables of phonetic stress will be discussed in §2.4.

## 2.2. Tone Languages

As stated above, tone languages mark each word or syllable with lexically contrastive pitch. For example, in Igbo, a Niger-Congo language spoken in Africa, each syllable is marked with either a high or a low tone. Disyllabic words have four possible tone patterns, demonstrated with the minimal pairs shown in (1).

(1) a. ákwá
'crying'
b. ákwà
'cloth'
c. àkwá
'egg'
d. àkwà
'bed' (Hymen, 1975:213)

(1a) shows a high-high prominence pattern, (1b) shows a high-low pattern, (1c) shows a low-high pattern and (1d) shows a low-low pattern.

Tones can also be *contour tones*, for example, rising or falling tones. Peking Mandarin is a tone language in which three of the four tones are contour. For example, (2a) has a high level tone, (2b) has a falling tone, (2c) has a rising tone and (2d) has a falling-rising tone.

(2) a. mā
'mother'
b. mà
'scold'
c. má
'hemp'
d. mǎ
'horse' (Hymen, 1975:214).

### 2.3. Phonological Stress: Pitch Accent vs. Stress Accent

Languages with phonological stress employ methods of marking some syllables more phonetically prominent than others. Sometimes this prominence is lexically contrastive, as in the Japanese examples listed in (3).

(3) a. káme
'turtle'
b. kamé
'jug' (Beckman 1988:146)

I will refer to this as *lexical prominence.*

In some languages, the assignment of prominence is derived through phonological rules. For example, English nouns and unsuffixed adjectives[2] assign phonological stress to the penultimate syllable if it is heavy. (In English, heavy syllables are those with long vowels or consonants in the coda.) This pattern can be seen in the words *agénda, tomáto* and *flamíngo*. If the penultimate syllable is light (an open syllable with a short vowel), then the antepenultimate syllable is assigned phonological stress. The words *álgebra, élephant* and *béautiful* demonstrate this pattern. I will refer to this type of phonological stress as *fixed prominence.* Prominence is fixed in the sense that phonological rules determine its placement. Every relevant word has the same prominence pattern.

Both lexical and fixed prominence can be manifested phonetically in one of two ways. The first way is through pitch accent. The Japanese words in (3) are an example of lexical prominence in a pitch accent language. According to Kaneko's analysis, she suggests that Blackfoot is an example of a pitch accent language with fixed prominence.[3] However, her assertion that Blackfoot is a pitch accent language has yet to be proven phonetically, which is precisely the topic of this thesis.

Pitch accent languages modulate $F_0$ to mark syllabic prominence. For example, in Tokyo Japanese, the prominent syllable has a higher pitch than the syllables around it (other dialects of Japanese use a low pitch to mark the prominent syllable) (Beckman 1986). The placement of a high pitch determines the pitch patterns of the whole word. The portion of a word in Tokyo Japanese before the prominent syllable has a mid level pitch, while the portion after the high pitch of the prominent syllable has a low pitch.

The second way prominence can be manifested is through phonetic stress. An example of lexical prominence manifested as phonetic stress can be seen in the Russian words in (4).

[2] In English, different word categories show different prominence patterns.
[3] Blackfoot also has lexical prominence in some forms. The overview of Blackfoot prominence patterns in §3.3 further discusses this.

(4) a. muká
'flour'
b. múka
'torture'
c. zamók
'lock'
b. zámok
'castle' (Dobrovolsky pers. comm.)

Fixed prominence manifested as phonetic stress can be seen in the English nouns and unsuffixed affixes discussed above. Some of the acoustic variables that are modulated in languages with phonetic stress have been claimed to be pitch, amplitude, duration, and vowel quality (Fry 1958; Beckman 1986; Dobrovolsky 1999; Sluijter and van Heuven 1996). These variables will be discussed further in §2.4.

Table 1 summarizes the four types of phonological stress, giving examples of languages in each category.

**Table 1**

| | Lexical Prominence | Fixed Prominence |
|---|---|---|
| Pitch Accent | Japanese | Blackfoot? |
| Phonetic Stress | Russian | English nouns, French, Chuvash |

## 2.4. Acoustic Correlates of Phonetic Stress and Pitch Accent

As we have seen, in tone and pitch accent systems, prominence can be related to $F_0$. Phonetic stress, on the other hand, has a number of acoustic correlates. Fry (1958) suggests that there are four main acoustic correlates to the perception of stress in English. The first is the duration of syllables. The second is the intensity (or amplitude) of the sound wave, which is related to the perception of loudness. The third acoustic correlate is the frequency of the sound wave, which is perceived as pitch. The fourth correlate is variations in the formant structure of the vowels, which is related to differences in sound qualities. For example, in English, an unstressed vowel is often reduced to a schwa. This can be demonstrated in the differences in the syllables of the noun *record*, /**ˈrɛ.kɔɹd**/, and the verb *record*, /**rə.ˈkɔɹd**/. Fry examined the first three of these physical correlates and determined that they all played a role in the listener's perception of prominence. He studied computer generated English noun-verb pairs listed in (5).

(5) a. ˈ**sʊb.jɛct** vs. **səb.ˈjɛct**
'subject:NOUN' 'subject:VERB'
b. ˈ**ab.dʒɛct** vs. **əb.ˈdʒɛct**
'object:NOUN' 'object:VERB'
c. ˈ**daj.dʒɛst** vs. **də.ˈdʒɛst**
'digest:NOUN' 'digest:VERB'
d. ˈ**kan.tɹækt** vs. **kən.ˈtɹækt**
'contract:NOUN' 'contract:VERB'
e. ˈ**pəɹ.mɪt** vs. **pəɹ.ˈmɪt**
'permit:NOUN' 'permit:VERB'

By varying the duration, intensity, and $F_0$ patterns he looked at which of these variables caused a shift in a listener's perception of prominence. He found that both duration and intensity played a role in the perception of prominence, although changes in intensity were not as strong of a variable because intensity changes alone never caused a complete shift in perception of prominence from one syllable to another. In both duration and intensity, the magnitude of the difference was important. When he looked at changes in $F_0$, he found that it also played a role in stress perception: the higher the frequency of the vowel in one syllable relative to the other, the more likely that syllable will be perceived as stressed. However, he also found that the magnitude of the frequency change was not important, it only mattered that a frequency change was present. Fry also examined listener's perceptions of words in which one syllable had a level pitch and the other syllable had a contour pitch. (Contour pitches included falling, rising, level-falling and level-rising. The contour pitches either began or ended at the same pitch as the syllable with a level pitch.) He found that the syllable with the contour pitch was perceived as prominent in two-thirds of the trials, and the level pitch was perceived as prominent in one-third of the trials. These results suggest that a change in pitch, regardless of the whether it is higher or lower, is an acoustic correlate of stress in English.

More recently, Beckman (1986) compared the acoustic correlates of Japanese pitch accent and English (phonetic) stress. Using Japanese and English as prototypes for pitch accent and phonetic stress respectively; she distinguished the acoustic differences between these two systems of marking prominence. Beckman used disyllabic minimal pairs in which the only difference was placement of the prominent syllable. For Japanese, she used the words listed in (6).

(6) a. íken
'opinion'

b. iken
   'differing view'
c. íkken
   '(one) house'
d. ikken
   'glance'
e. kábu
   'lower part'
f. kabu
   'stocks'
g. káme
   'turtle'
h. kamé
   'jug'
i. kámi
   'god'
j. kamí
   'paper
k. káta
   'shoulder'
l. katá
   'form' (Beckman 1986:146)

For English, she used the same noun-verb pairs used by Fry listed above in (5).

Words were recorded in carrier sentences from native speakers of each language. The carrier sentence for the Japanese words is shown in (7).

(7) Sosite _______ to iimasu.
    'Next I'll say _______' (Beckman 1986:145)

The carrier sentence for the English words is shown in (8).

(8) I said _______ this time. (Beckman 1986:146)

The purpose of the carrier sentence was to control for the intonation so that it would be the same for each word analyzed.

The words were then analyzed according to five variables. For each vowel she measured the $F_0$, the duration, the average amplitude (in dB), the amplitude peak, and the total amplitude. $F_0$ values were taken by measuring the third harmonic on narrow band spectrograms, and dividing the value by three. The measurement was taken at the obvious peak if there was one, if not, the measurement was taken midway through the syllable nucleus. The duration was measured on a wide band spectrogram using spectral cues such as the burst spike at the end of stop closures. The amplitude peak was the highest point of energy in dB within the vowel. The total amplitude was obtained by adding up all of the amplitude values for the vowel taken at ten millisecond intervals. This measurement takes into account both amplitude and duration. The last measurement, average amplitude, took the total amplitude values, divided by the duration, and multiplied by ten (for the ten millisecond intervals).

Beckman found that in Japanese, a prototypical pitch accent language, the only significant variable for marking prominence was the $F_0$. However, In English, a prototypical phonetic stress language, she found that all the variables measured were significant acoustic correlates of phonetic stress. The most significant variable was total amplitude, which is a reflection of both duration and intensity.

Dobrovolsky (1999) examined the acoustic correlates of prominence in Chuvash, a Turkic language spoken in Russia. Chuvash is a language with fixed prominence. Prominence is placed on the last full vowel of a word. If there are no full vowels (that is, if a word contains only the reduced/central vowels ĕ and ă), there is no phonological stress (Dobrovolsky 1999). The data in (9) show this pattern.

(9)
a. á.dăl
'Volga'
b. i.kér.čĕ
'pancake'
c. šu.paš.kár
'Cheboksary'
d. lĕ.bĕs
'butterfly'
e. pĕ.lĕ.vĕ.mĕr
'we knew' (Dobrovolsky pers. comm.)

Dobrovolsky (1999) examined disyllabic words, comparing four word groups: words with a full vowel in the first syllable and a full vowel in the second syllable (full-full), full-reduced words, reduced-full words and reduced-reduced words. He looked at the same variables studied by Beckman to determine whether Chuvash patterned with Japanese as a pitch accent language or English as a language with phonetic stress.[4] He found that the $F_0$ variable was not a significant correlate of prominence in Chuvash. There were significant differences in duration and total amplitude. This suggests that Chuvash has phonetic stress, in accord with Beckman's analysis of English.

Another acoustic analysis of a language with phonetic stress was done by Sluijter and van Heuven (1996). They studied the acoustic correlates of prominence in Dutch, a language similar to English in which phonetic stress can be lexically contrastive. They examined the Dutch minimal pair shown in (10).

(10) a. ˈka:nɔn
canon
'cannon'
b. ka:ˈnɔn
kanon
'canon' (Sluijter and van Heuven 1996:2473)

They also used nonsense words in which each syllable of the target word (from (10) above) was replaced by *na*. In an attempt to separate the phonetic stress from intonation, they put each target and each nonsense word in different carrier phrases that placed the sentence intonational prominence on the target word in one instance and on a different word in another instance. They found that $F_0$ is not a correlate of word prominence in Dutch, but instead is a correlate of intonational prominence within a phrase.

Sluijter and van Heuven also looked at duration, vowel quality, overall intensity (in decibels) and intensity at different frequency levels. They found that the most reliable correlate of prominence in Dutch is duration, and the next most reliable correlate was what they called spectral balance, which is an increase in the intensity in the higher frequencies – above 0.5 kHz. Vowel quality and overall intensity turned out to be poor cues to phonetic stress in Dutch. Sluijter and van Heuven state that overall intensity is related to the intonational prominence of phrases, along with $F_0$. They suggest that this is the reason that earlier studies such as Beckman's found overall intensity to be a correlate of phonetic stress.

The present study will follow those of Beckman (1986) and Dobrovolsky (1999) and will examine $F_0$, duration, amplitude peak, average amplitude and total amplitude as possible correlates of prominence in Blackfoot.

[4] This was not the only purpose of this study. He also looked at whether Chuvash has left-edge default phonological and phonetic stress, that is, stress falling on the first vowel or syllable of a word with only reduced vowels, as had been claimed in earlier studies. He claimed that words with a reduced-reduced pattern did not have stress and that the apparent left-edge default stress was a result of falling word or phrase intonation found on the majority of words, irrespective of stress type.

## 3. An Introduction to Blackfoot[5]

### 3.1. Phonemic Inventory[6]

Modern Blackfoot contrasts five places of consonant articulation: bilabial, alveolar, palatal, velar and laryngeal. It also contrasts five manners of articulation: plosives, affricates, fricatives nasals and glides. The proposed phonemic inventory of the Modern Blackfoot consonants is shown in Table 2.

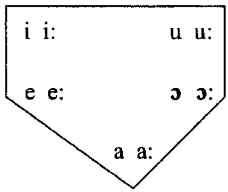
**Table 2: Phonemic Inventory of Blackfoot Consonants**

| | Labial | Alveolar | Palatal | Velar | Glottal |
|---|---|---|---|---|---|
| Plosives | p | t | | k | ʔ |
| Affricates | ps | ts | | ks | |
| Fricatives | s | | | x | |
| Nasals | m | n | | | |
| Glides | w | | j | | |

Note that these consonants also have phonemic geminate counterparts.

There are five vowels in the Modern Blackfoot phonemic inventory, shown in Table 3.

**Table 3: Phonemic inventory of Blackfoot vowels**




These vowels also have contrasting long counterparts. Blackfoot also has one diphthong, /oj/, but see the discussion below on the orthography for evidence of other diphthongs in other dialects of Blackfoot.

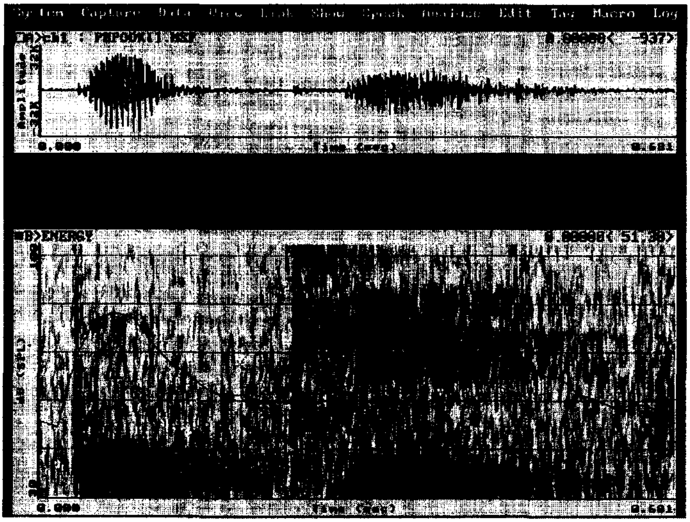
---

[5] For a discussion of Blackfoot Syntax, see Appendix A.

[6] For a discussion of Blackfoot phonology, see Appendix B.

Vowels are devoiced word-finally in Blackfoot. During elicitations, it seemed as if only some vowels are devoiced word-finally. However, in words that had a word-final voiced vowel, spectrographic analyses revealed that the latter half of the vowel was still devoiced. This can be seen in Figure 3-1. The line indicates intensity of the sound wave in decibels. This process of word-final devoicing may extend to the whole syllable (as discussed in §3.1). Vowels are also devoiced word-internally before [x]. Just as word-final devoicing, either the whole vowel or the first half of the vowel is devoiced. It may be that vowel devoicing is mandatory word-finally and before [x], and the cases where only half of the vowel is devoiced are underlying long vowels. This is in fact how Frantz (1991:6) orthographically represents vowels before [x]. If a vowel is voiced before [x] it is long, if it is voiceless, it is short. I propose that this analysis would be the same for word-final devoicing, except that vowel devoicing can extend to the first half of a long vowel (in fact, the whole syllable) in free variation.

**Figure 1:** ***pókii*****, 'small' (PB)**

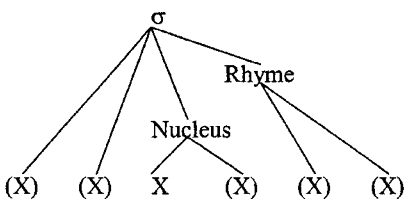


A Roman orthography was developed based on the phonological inventory of Blackfoot. Most of the consonants are written the same as their English counterparts. However, there are two phonemes that are not represented in English orthography, the velar fricative /x/ and the laryngeal stop /ʔ/. These are written as *h* and ' respectively. The three vowels, /i/, /u/ and /a/, are written with the corresponding Roman orthography *i, u* and *a*. The other two vowels, /e/ and /ɔ/, are written as the diphthongs *ai* and *ao*. Frantz (1991) suggests that some varieties of Blackfoot still pronounce these as diphthongs (from which they were historically derived). However, our elicitations showed that they were produced as monophthongs in the variety spoken by our consultants. Kaneko (1999:12-16) analyzes these sounds as diphthongs and thus has a tripartite vowel inventory with just /i/, /u/ and /a/. The consultants used in her study were from the Blood reserve, so this is likely a difference in dialect. Prominence is shown in the orthography by an acute accent over the vowel of the prominent syllable.

### 3.2. Syllable Structure

The simplest syllable in Blackfoot is a syllable with a short vowel and no onset or coda. Blackfoot also allows long vowels and diphthongs. Onsets and codas are also optional in Blackfoot, which allows for both simple and complex onsets and codas. Words are syllabified according to the Maximum Onset Principle, which states that intervocalic consonants and consonant clusters become onsets rather than codas as long as the consonant cluster is an allowable sequence within the language. Consonant clusters allowed in Blackfoot involve the alveolar fricative before or after stops and affricates. As earlier stated, Blackfoot also has geminate consonants, and these can also occur as part of a consonant cluster. When syllabified, geminates will generally become both the coda of the preceding syllable and the onset of the following syllable. The Maximal Syllable Template in Blackfoot is shown in (11).

(11) *Blackfoot Maximal Syllable Template (after Kaneko, 1999)*[7]




[7] This is the syllable template given by Kaneko (1999). It may be that the template should include a node for an onset. However, this issue is not relevant for the current analysis and will therefore not be discussed.

### 3.3. Blackfoot Prominence Patterns[8]

Kaneko did an analysis of prominence patterns in Blackfoot nouns. She suggests that Blackfoot has a mixed lexical and fixed accent system. Kaneko does not explain how the placement of the pitch accent was determined for her study. Presumably it was a subjective analysis based on what she heard during elicitations with her consultants.

Just as in Chuvash and English, Blackfoot phonological stress patterns depend on syllable weight. In Blackfoot, heavy syllables are those with a long vowel or a consonant in the coda. On mono-morphemic nouns, a heavy syllable carries prominence. This can be seen in (12a). If there is no heavy syllable, prominence is lexically specified, as in (12b) and (12c). If there is more than one heavy syllable, prominence will fall on the leftmost heavy syllable, as seen in (12d). There do not appear to be any phonological rules to predict where the prominence will fall in words with only light syllables.

| | | | |
|---|---|---|---|
| (12) | a. /ˈnaa.pi/ | náápi | 'trickster' |
| | b. /ˈsto.ʔo/ | stó'o | 'ghost' |
| | c. /so.ˈpo/ | sopó | 'wind' |
| | d. /piˈksii.ksii.na/ | piksíiksiina | 'snake' (Kaneko 1999:85-87,145) |

The following patterns are proposed by Kaneko for compound nouns. The leftmost accent of the two morphemes is retained, unless the accent is on the final syllable, in which the accent moves to the juncture between the two morphemes. This can be seen in (13a) and (13b) respectively. If both morphemes of the compound noun have no lexical accent, the final syllable is marked as prominent. This is shown in (13c).

| | | | |
|---|---|---|---|
| (13) | a. /piˈsat-/ + /ˈna.pi.nju.wan/ | pisát- + nápinyowan | 'fancy/unusual' + sugar' |
| | /piˈsa.tsa.pii.nju.wan/ | psátsaapiinyoan | 'candy' |
| | b. /sik/ + /ox.ˈkii/ | sik- + ohkíí | 'black' + 'water' |
| | /si.ˈkox.kii/ | sikóhkii | 'vanilla extract' |
| | c. /sik-/ + /-i.ka/ | sik- + -ika | 'black' + foot' |
| | /si.ksi.ˈkaa/ | Siksika | 'Blackfoot' |
| | | | (Kaneko 1999:153, 155, 156) |

[8] Note that this is only a brief summary of Kaneko's (1999) analysis. In her paper, she lists many more examples and constraints than will be discussed here. See her work for a more detailed look at prominence patterns in Blackfoot nominals.

Kaneko also notes that there is a contrast found in long vowels; those with a high level pitch, those with a rising pitch and those with a falling pitch. This suggests that both timing slots of a long vowel can carry prominence, or just one can carry prominence. She claims that the words in (14) show these differences.

| | | | |
|---|---|---|---|
| (14) | a. [pií.ta] | piíta | 'eagle' |
| | b. [póo.ka] | póoka | 'child' |
| | c. [aa.kíí] | aakíí | 'woman' (Kaneko 1999: 142-143) |

(14a) shows a rising pitch in the first syllable. The first syllable of (14b) shows a falling pitch. The second syllable in (14c) shows a high level pitch. In addition to examining the acoustic correlates of Blackfoot prominence, this study will test Kaneko's claim of different prominence patterns on long vowels by examining spectrograms for the above patterns of prominence. The results can be found in §6.1.1.

## 4. Acoustic Analysis

### 4.1. Procedures

The present study assesses Blackfoot phonetic accent/prominence with the same five variables calculated by Beckman (1986). The first is $F_0$, which is measured in Hertz (Hz). The second and third variables are the average amplitude and amplitude peak. Amplitude is measured in decibels, which is related to the perception of loudness. The fourth variable is total amplitude. This is a measurement of decibels over time. The decibel level is measured at 20ms intervals over the course of a vowel, then all of the measurements are totaled. Total amplitude thus takes into account both the amplitude and the duration of the vowel. The final variable is duration. This is measured in milliseconds (ms).

The results will also be expressed in ratios. These ratios express the results of the measurements from each word with one figure, rather than having a value for each vowel. The formulas in (15) were used to calculate the ratios.

(15) a. $F_0$ Peak Ratio = 17.31 ln (Hz $V_2$/Hz $V_1$)[9]
b. Duration Ratio = ln (ms $V_2$/ms $V_1$)
c. Amplitude Peak Ratio = dB Peak $V_2$ – dB Peak $V_1$
d. Average Amplitude Ratio = dB Average $V_2$ – dB Average $V_1$
e. Total Amplitude Ratio = Total Amplitude $V_2$ – Total Amplitude $V_1$

[9] This formula is used by Beckman to calculate the difference in semitones between the first and second syllable.

With these formulas, a word with a higher value of the variable being measured in the first vowel as compared to the second (falling values) will result in a negative ratio value and words with a higher value in the second vowel (rising values) will result in a positive ratio value.

Most of the data used for this study was collected over the course of four months as described in §1.2. Disyllabic words from the corpora were selected for the present study. The only disyllabic words not used were those with a glide between the vowels of the first and second syllable,[10] and words in which there was too much background noise to accurately interpret the spectrograms.

Thirty-eight words were used, 24 from speaker PB and fourteen from speaker NB. The words were elicited in isolation. The consultants were asked off-tape what the target word was, and then they were asked to repeat the word twice while being recorded. For this analysis, the first of the two words was analyzed, although there were a few instances where the word was only recorded once.

For the $F_0$, average amplitude and amplitude peak analyses, the words were divided into two groups. The first group consisted of words with a strong-weak prominence pattern and the second group consisted of words with a weak-strong prominence pattern. The duration and total amplitude measurements needed to be further divided because Blackfoot has a phonemic distinction between long and short vowels. For these two measurements, each group was subdivided into four more groups: words with a short vowel followed by a short vowel, words with a short vowel followed by a long vowel, words with a long vowel followed by a short vowel, and words with a long vowel followed by another long vowel. Words with different prominence patterns could then be compared between words with the same patterns of vowel length. The total number of tokens for each group can be seen in table 4.

**Table 4: Number of tokens per group**

| | Strong-weak | Weak-strong |
|---|---|---|
| Short-short | 10 | 6 |
| Short-long | 5 | 5 |
| Long-short | 3 | 2 |
| Long-long | 2 | 5 |

Comparisons will only be made in groups with a short-short or short-long length pattern due to the small number of words in the other two groups.

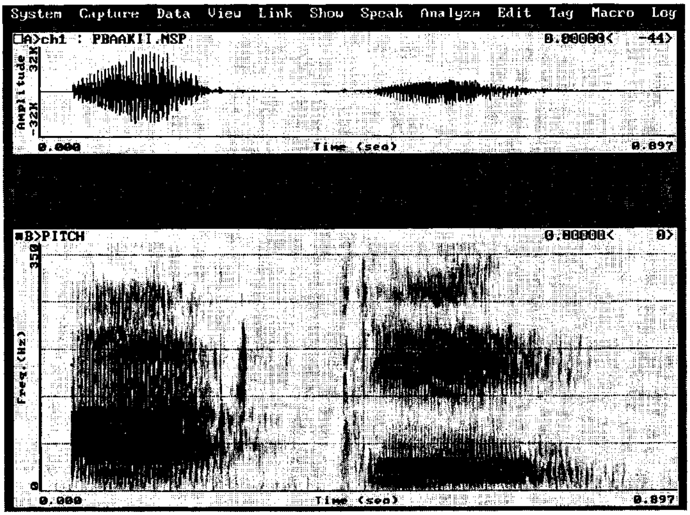
[10] The forms were eliminated because it was too difficult to determine where the first vowel of the word ends and the second vowel begins.

Vowel qualities are randomly distributed in the data. The three most common vowels, /i/, /o/ and /a/, occur in both the first and second syllable in words with a strong-weak and weak-strong prominence pattern, with the exception of /o/ in strong-weak words.

The data was analyzed on a Kay CSL model 4300, version 4.17. The words were digitized from the cassette tape at a sampling rate of 12800 Hz. Each word was examined on a broadband spectrogram with a bandwidth of 188 Hz.

The $F_0$ was measured by using a peak-picking algorithm that sampled the $F_0$ at 20 ms intervals. Figure 2 shows these pitch patterns overlaid on the broadband spectrogram. The dots are pitch peaks in Hz. The waveform is in the window above the spectrogram.

**Figure 2:** ***aakíí*, 'woman' (PB)**




In order to determine the accuracy of this algorithm, random words were checked by measuring the fifth harmonic on a narrowband spectrogram with a bandwidth of 31 Hz and dividing by five. These measurements agreed with those given by the algorithm. Figure 3 shows the wideband spectrogram of the same word in Figure 2. The arrow is pointing to the fifth harmonic. The amplitude peak and average amplitude were determined by using a peak-picking algorithm that calculated the dB level at 20 ms intervals. The amplitude can be seen in Figure 4.

**Figure 3:** ***aaklí*, 'woman' (PB)**

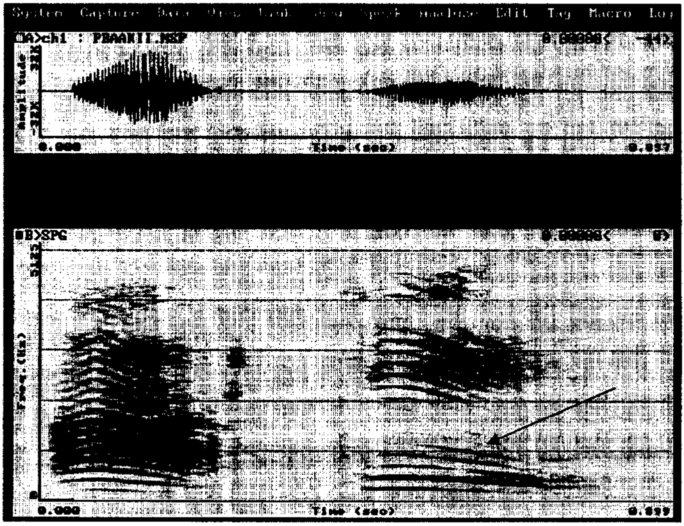
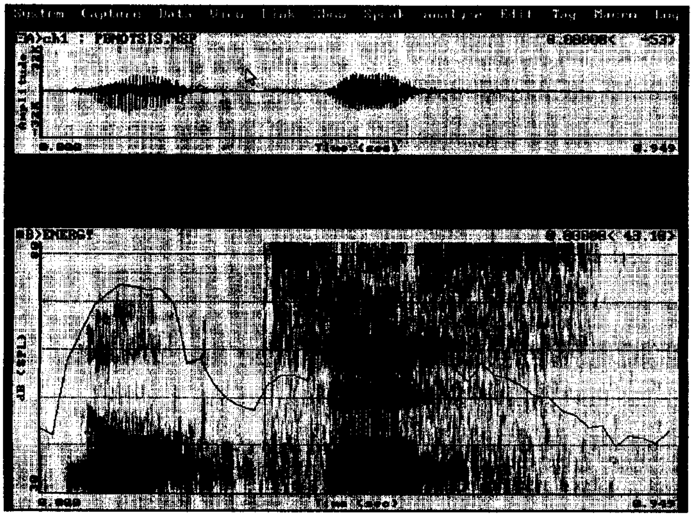


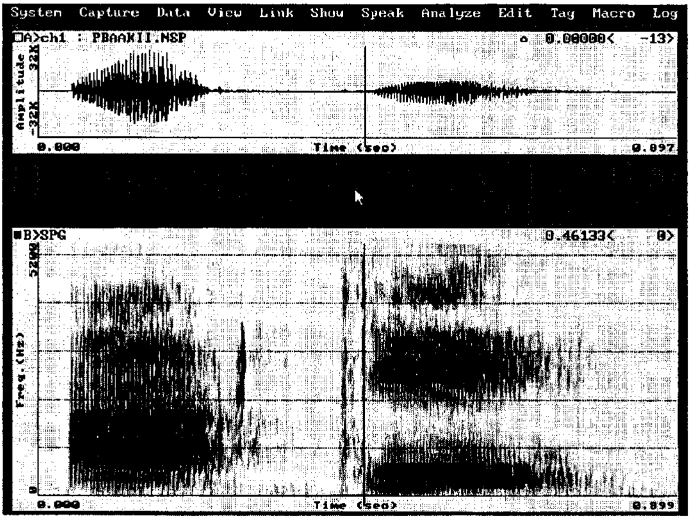
**Figure 4:** ***notsís*, 'my arm/hand' (PB)**



For the average amplitude, all decibel values during a value were added and then divided by the number of measurements taken. The total amplitude was the value obtained in the calculation of the average amplitude before dividing.

The duration measurement was the time in ms of thc vowel. Duration measurements were made on the broadband spectrogram. The measurement of the vowel began after the burst spike of the preceding stop if it was an obstruent and after the end of the nasal murmur. This can be seen in figures 5 and 6. The vertical line in each figure indicates where the duration measurements began.

**Figure 5: *aakíí*, 'woman' (PB)**




The measurement of the end of the vowel duration was more difficult to determine. Generally, vowels end before a portion where there is no sound production (unless followed by a nasal), either as a result of the closed portion of a stop or at the end of the word. This was especially difficult when the latter half of the vowel was voiceless. Due to recording limitations, the intensity of the vowel formants is, at times, the same intensity of the background noise. Measurements were taken at the point where the formants were no longer apparent. This can be seen in Figure 7.

**Figure 6:** ***ninàá*, 'man/chief' (PB)**

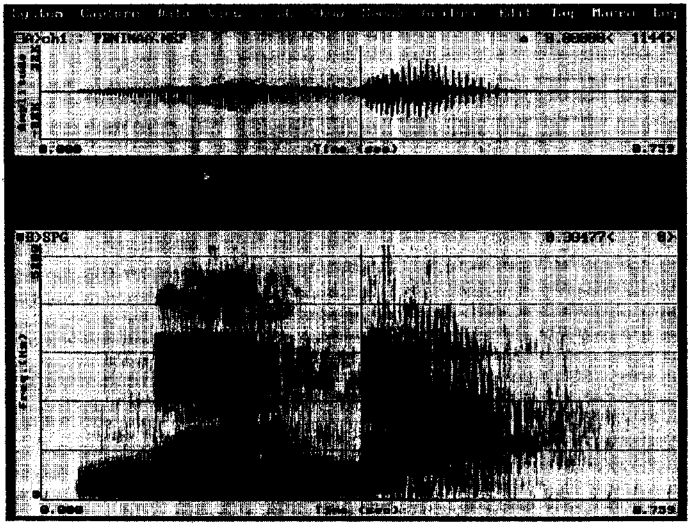
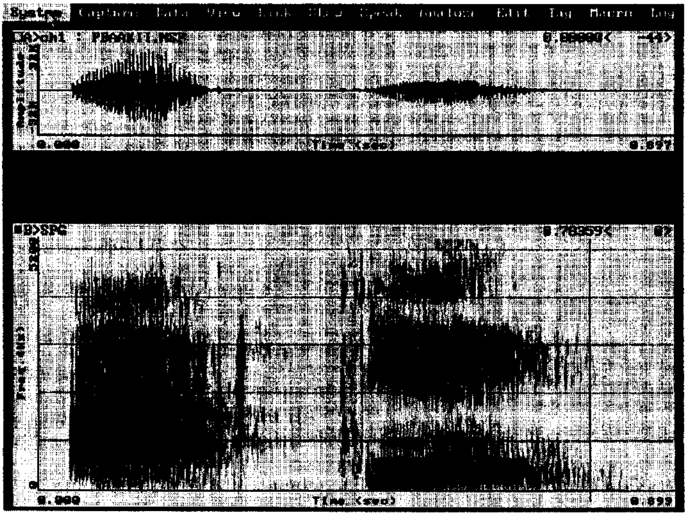


**Figure 7:** ***áákii*, 'woman' (PB)**



## 4.2 Results

This section gives the results of the acoustic analysis. The different measurements for each vowel will be presented as well as the ratios discussed in §4.1. The raw data will be presented in the following sections, while the statistical analyses will be presented in Chapter 5. This will determine whether or not any differences found in the raw data are statistically significant.

The mean $F_0$ for the first vowel of words with a strong-weak prominence pattern is 201.30 Hz, and the mean for the second vowel is 158.15 Hz. This shows a definite drop in pitch from the first vowel to the second. Figure 8 shows this pattern. The mean $F_0$ for the first vowel of words with a weak-strong prominence pattern is 187.66 Hz, and the mean for the second vowel is 189.16 Hz. There does not seem to be a considerable difference in pitch between the two vowels in words with a weak-strong prominence pattern. Figure 9 shows this pattern.

The level $F_0$ pattern may be a result of a falling intonation overlaid on the prominence patterns of the word. A falling intonation plus a rising prominence pattern results in a general flattening of pitch from one vowel to the next. In words with a strong-weak prominence pattern, the falling pitch is augmented by the falling intonation.

**Figure 8:** ***píítaa*****, 'eagle' (PB)**

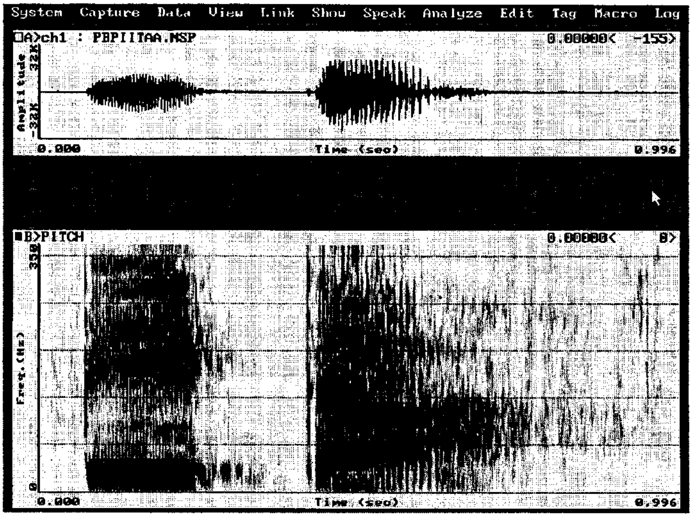
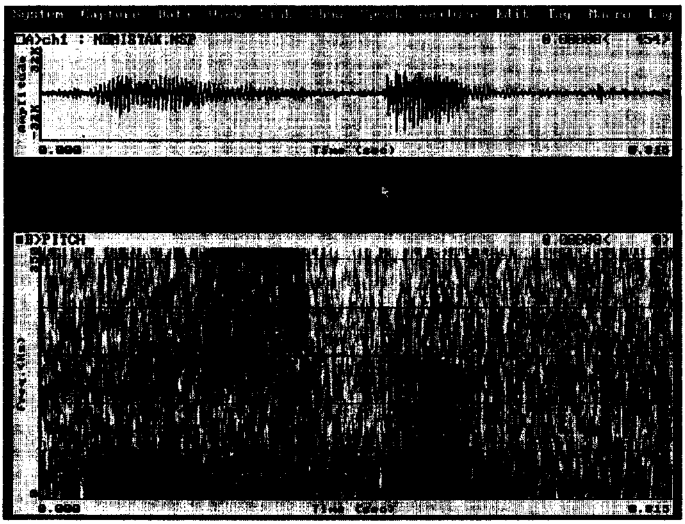
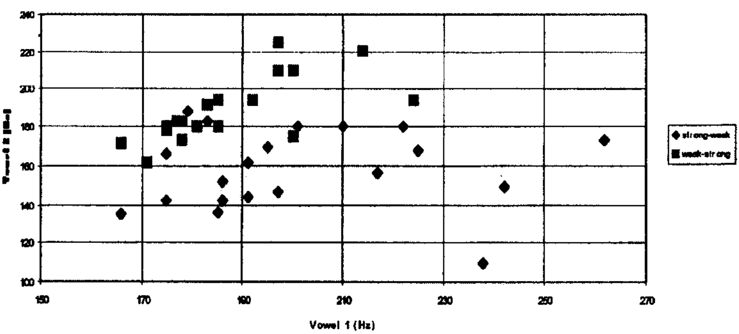


**Figure 9: *misták*, 'mountain' (NB)**



The results of the $F_0$ analysis are shown in the scatter plot in table 5.[11]

**Table 5: $F_0$ Peak**




[11] All graphs were produced with Microsoft Excel 2000.

This chart plots the $F_0$ peak of the first vowel against the second vowel. Words with a higher $F_0$ peak in the second vowel as compared with the first vowel will be closer to the top left corner of the scatter plot. You can see that the words examined follow this general pattern, with most of the words with a weak-strong prominence pattern falling to the upper left of the words with a strong-weak prominence pattern.

The average $F_0$ Peak Ratio for words with a strong-weak prominence pattern is -4.192 and the average for words with a weak-strong prominence pattern is 0.121. The negative ratio value indicates a falling pitch while the positive ratio value indicates a rising pitch. This demonstrates that words with a higher pitch on the first vowel correspond to words with a strong-weak prominence pattern and words with a higher pitch on the second vowel correspond to words with a weak-strong prominence pattern. The ratio value for the weak-strong words is close to zero. This reflects the fact that there was not a large difference in the $F_0$ of the two vowels. These results suggest that pitch is a correlate of Blackfoot prominence.

Because Blackfoot distinguishes between long and short vowels, duration measurements could not be compared between all of the words with a strong-weak prominence pattern and all of the words with a weak-strong prominence pattern. Table 6 shows the words with a short-short length pattern and Table 7 shows the words with a short-long length pattern. Words with a longer duration on the first vowel as compared to the second vowel will appear closer to the top left portion of the charts.

**Figure 6: Duration in Short-Short Words**

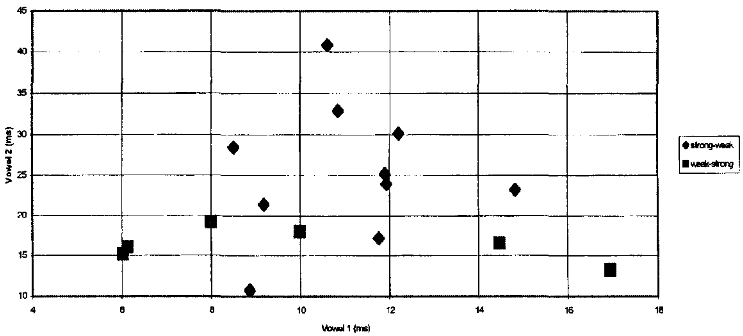
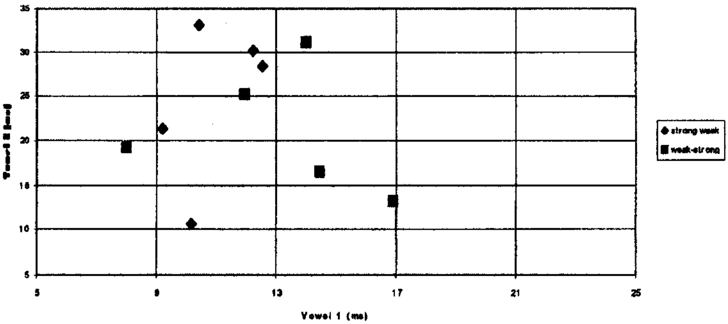


**Table 7: Duration in Short-Long Words**




The pattern is not as clear in duration as it seemed to be for $F_0$. The mean length of the first vowel in short-short words with a strong-weak pattern is 11.061 ms while the mean length of the second vowel is 25.392 ms. In short-short words with a weak-strong prominence pattern the mean length of the first vowel is 10.257 ms and the mean length of the second vowel is 16.376 ms. There is a definite lengthening of the final vowel, which is probably due to the intonational properties of phrase-final lengthening. However, the final vowel in words with a strong-weak prominence pattern is lengthened more than in words with a weak-strong prominence pattern. This can be seen in figures 10 and 11 respectively. Beckman found the prominent syllable is lengthened in English. The results of Blackfoot prominence patterns show the opposite is happening. The prominent syllable is shorter.

Figure 10: *nitán*, 'my daughter' (PB)

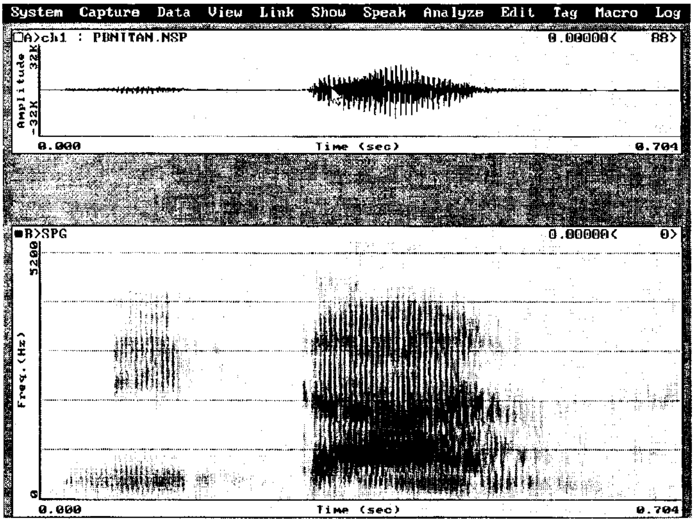
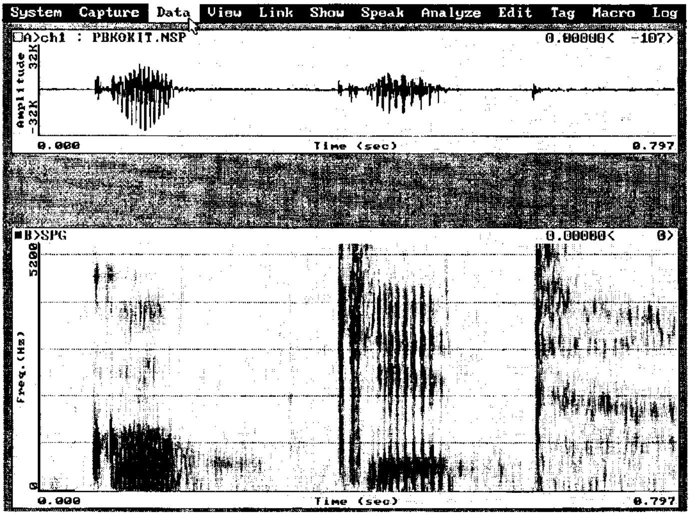



Figure 11: *kókit*, 'give' (PB)



In words with a short-long length pattern, the mean duration of the first vowel of those with a strong-weak prominence pattern is 9.557 ms and the mean duration of the second vowel is 32.149 ms. In short-long words with a weak-strong pattern, the mean duration of the first vowel is 11.384 ms and the second vowel is 28.432 ms. Again, the second vowel is lengthened more in those words with a strong-weak pattern, which points to the fact that in Blackfoot, duration does not pattern in the same ways it does in languages with phonetic stress.

The average duration ratio in short-short words with a strong-weak prominence pattern is 0.597 and the average for words with a weak-strong prominence pattern is 0.623. The positive number for both word types reflects the fact that the final vowel of a word was always lengthened in the elicitations. The value is higher in words with a weak-strong prominence pattern than those with a strong-weak prominence pattern, which is what we would expect if duration were a correlate of Blackfoot prominence. A higher positive ratio value indicates a greater lengthening in the final syllable.

In short-long words with a strong-weak prominence pattern, the average duration ratio is 1.026 and the average in words with a weak-strong prominence pattern is 0.957. Contrary to the short-short words, the average duration ratio for words with a strong-weak prominence pattern is higher than the average for words with a weak-strong pattern. This is not the expected result if duration were a correlate of Blackfoot prominence.

Due to the fact that there is no clear pattern in Tables 6 and 7, and the fact that short-short and short-long words show different patterns in the duration ratio, duration is probably not a correlate of Blackfoot prominence.

The mean amplitude peak value for the first vowel in strong-weak words is 77.061 dB, and the mean for the second vowel is 74.549 dB. Figure 12 shows an example of this pattern.

In words with a weak-strong prominence pattern, the mean amplitude peak value in the first value is 75.663 dB and 77.415 in the second vowel. An example of this pattern can be seen in Figure 13. These patterns do suggest that the prominent syllable in Blackfoot has a higher amplitude peak.

**Figure 12:** ***kókit*****, 'give' (PB)**

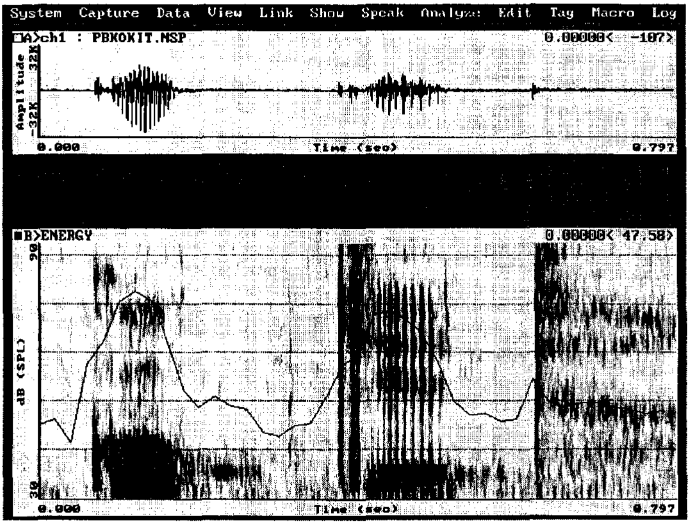



**Figure 13:** ***notsís*****, 'my arm/hand' (PB)**

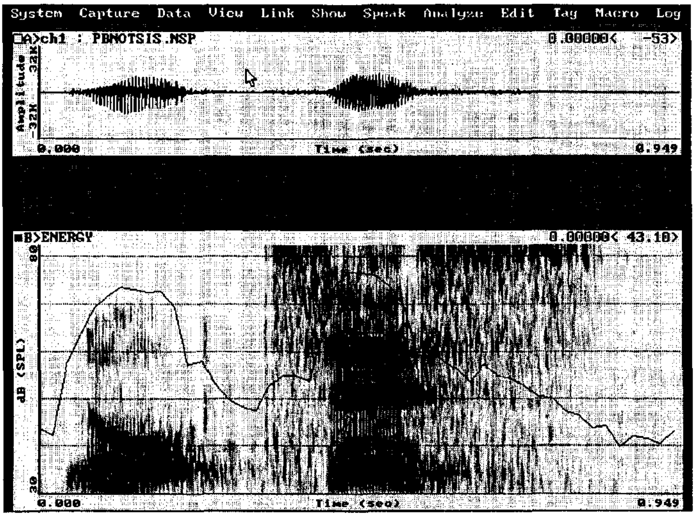
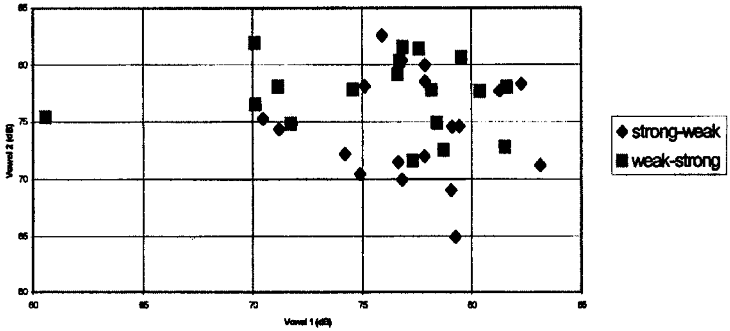


Table 8 shows the scatter plot for the amplitude peak.

**Table 8: Amplitude Peak**




Words with a higher amplitude peak in the second vowel will appear closer to the top left of the chart in table 8. The pattern is not as clear as the pattern in table 5 for $F_0$; the statistical analysis in section five discusses whether or not the difference is significant.

The mean amplitude peak ratio for words with a strong-weak prominence pattern is -2.507 and the average for words with a weak-strong prominence pattern is 1.752. This is what we would expect if amplitude peak is a correlate of Blackfoot pitch accent. Words with a higher amplitude peak on the first vowel correspond to words with a strong-weak (falling) prominence pattern and words with a higher amplitude peak on the second vowel correspond to words with a weak-strong (rising) prominence pattern.

In words with a strong-weak prominence pattern, the mean average amplitude value for the first vowel is 73.149 dB and the mean value for the second vowel is 63.844 dB. An example of this can be seen in figure 14.

In words with a weak-strong prominence pattern, the mean value for the first vowel is 70.587 dB and the mean value for the second vowel is 70.240 dB. Figure 15 shows an example of this pattern. This pattern is very similar to the pattern we saw in the $F_0$ values with a sharp drop in words with a strong-weak pattern and not much change in words with a weak-strong pattern. This could be due to the intonation overlay.

Figure 14: *pókii*, 'small' (PB)

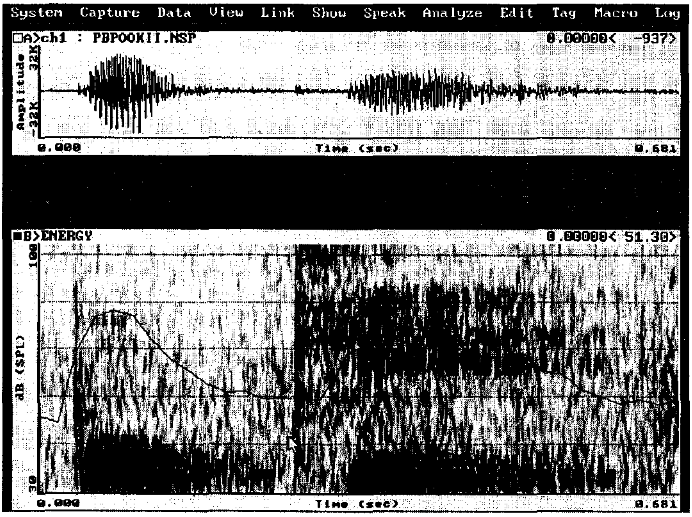



Figure 15: *pookáá*, 'child' (PB)

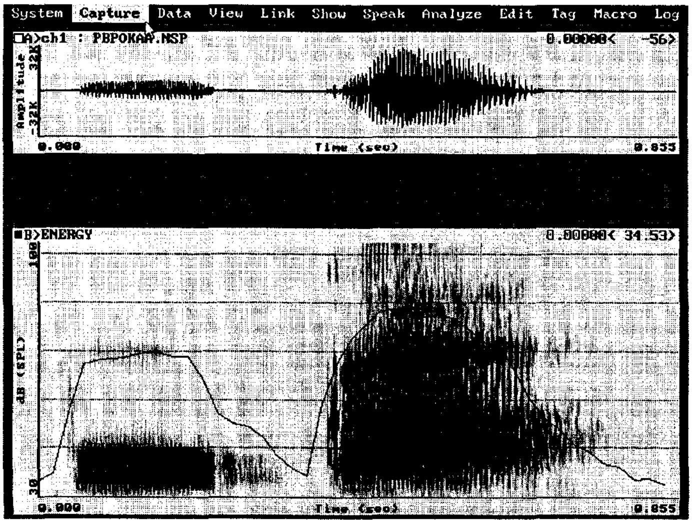
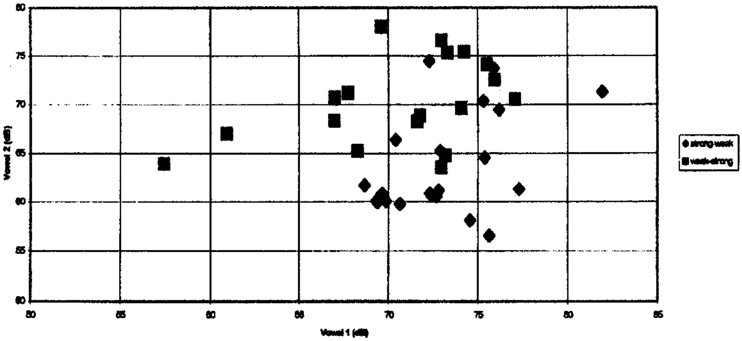


Table 9 shows the scatter plot of average amplitude values in vowel one versus vowel two.

**Table 9: Average Amplitude**

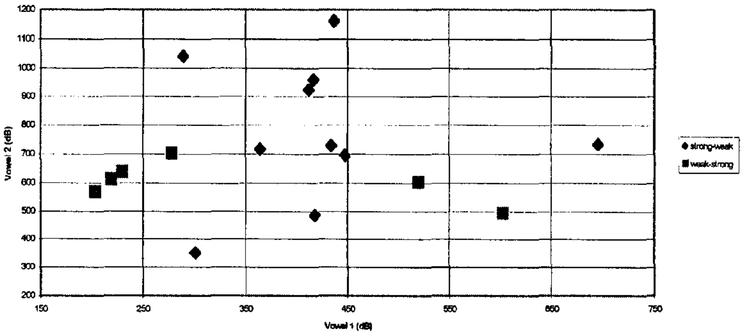



Words with a higher average amplitude value in the second vowel will be closer to the top left corner of the chart. There is a somewhat overlapping pattern of weak-strong words closer to the top left and strong-weak words closer to the bottom right.

The mean average amplitude ratio for words with a strong-weak prominence pattern is -9.292 and the average for words with a weak-strong prominence pattern is -0.345. The negative value in both prominence patterns indicates that the peak amplitude of the majority of the words elicited occurs during the first vowel. However, there is a considerable difference between the two values that corresponds to what we would expect if average amplitude is a correlate of Blackfoot pitch accent. Words with a higher average amplitude value on the first vowel correspond to words with a strong-weak prominence pattern and words with a fairly level average amplitude on the second vowel correspond to words with a weak-strong prominence pattern. The raw data suggests that average amplitude is a correlate of Blackfoot prominence.

Like duration, total amplitude values need to be looked at separately for each of the different length patterns. Words with a short-short length pattern and a strong-weak prominence pattern have a mean total amplitude of 421.447 dB for the first vowel and 781.173 dB for the second vowel. Short-short words with a weak-strong prominence pattern have a mean total amplitude of 342.116 dB for the first vowel and 603.870 dB for the second. Because total amplitude depends on duration, the lengthening of the final syllable that we saw in §4.2 will result in total amplitude values higher on the second vowel than the first. Like duration, this shows a pattern opposite to that found in English. Table 10 shows a scatter plot of these words.

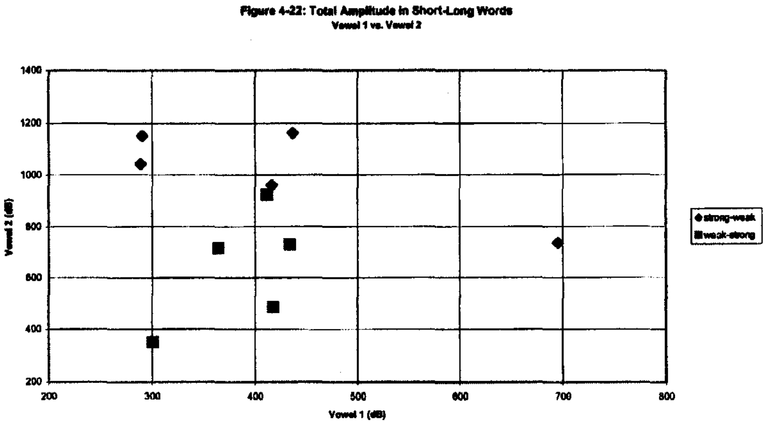
**Table 10: Total Amplitude in Short-Short Words**




In table 10, words with greater total amplitude on the second vowel will appear closer to the top left corner of the chart. Insofar as this chart does not seem to show any clear pattern, short-short words do not seem to be distinguished by total amplitude.

In words with a short-long length pattern and a strong-weak prominence pattern, the mean total amplitude value for the first vowel is 358.810 dB and 1101.564 dB for the second vowel. In short-long words with a weak-strong prominence pattern, the mean total amplitude value for the first vowel is 428.100 dB and 984.402 dB for the second vowel. Like the words with a short-short length pattern, these words seem to show that total amplitude does not pattern with prominence. Table 11 shows a scatter plot of the total amplitude in words with a short-long length pattern.

## Table 11: Total Amplitude in Short-Long Words




As in table 10, words with greater total amplitude on the second vowel will appear closer to the top left corner of the chart. This chart seems to show the opposite pattern than one would expect in a language with phonetic stress. The total amplitude values pattern in much the same way as the duration values. This is doubtless due to the fact that duration is an important component of the total amplitude measurement. The total amplitude measurement is greater when the duration of the vowel is longer.

The average total amplitude ratio in short-short words with a strong-weak prominence pattern is 359.726 and the average for words with a weak-strong prominence pattern is 261.753. As with the values for duration ratios, the positive number for both word types reflects the fact that the final vowel of a word was always lengthened in the elicitations. The value is higher in words with a strong-weak prominence pattern than those with a weak-strong prominence pattern.

In short-long words with a strong-weak prominence pattern, the average total amplitude ratio is 742.754 and the average in words with a weak-strong prominence pattern is 556.302. Like the short-short words, the value is higher in words with a strong-weak prominence pattern than those with a weak-strong prominence pattern. This is not what we would expect if total amplitude were a correlate of Blackfoot prominence.

The next section will look at whether or not the patterns found above are statistically significant.

## 5. Statistical Analyses[12]

### 5.1. Statistical Procedures

The statistical tests used in this paper to determine the significance of the results will be discussed in this section. Statistics will be used to determine whether or not there is a significant difference in each of the five variables between the words with differing prominence patterns.

Some terms used in this section will need defining. *Sample* refers to the set of data actually observed. In this paper, this will be the data obtained in the present experiments. The *population* refers to the complete set of data. For example, for the $F_0$ variable measured in this study, population would refer to the $F_0$ of the vowels in all disyllabic words ever uttered by any Blackfoot speaker of the same dialect being studied. Thus, the sample is a subset of the population. Obviously the population is not something that can be measured, so the statistics used in this paper will be used to determine what the probability that the difference found in the sample data is actually a reflection of a difference that would be found in the population data. An *independent variable* is a variable whose selection is decided on or controlled by the experimenter with the purpose of examining their effects on other variables. In this study the independent variables are prominence and length patterns of the words. A *dependent variable* is the variable being measured. This study has five dependent variables: $F_0$, duration, amplitude peak, average amplitude and total amplitude.

The *null hypothesis* ($H_0$) is the hypothesis that is being tested by the particular statistical procedure. Scientific reasoning suggests that although we can never prove any statement about a population to be true, we can prove it to be false, (assuming we do not have access to the whole population). This is because even if we observe all of the members of a sample to have a particular property, we still cannot conclude with absolute certainty that every member of the population has that property. This is true no matter how many members of the population are observed, short of observing all members. However, it takes only one member of the population without the property to prove that a statement about the population is false. This is the basic idea behind the $H_0$, which is the hypothesis that there is no difference (or no relationship) between the independent and dependent variables. In this study, the null hypothesis is that the values of the

[12] The statistical procedures used in this study were learned in Psychology 312 – Experimental Methods and Design in the Fall/Winter session of 1998/1999 at the University of Calgary. Dr. Christopher Sears, one of the course instructors, also provided help with the statistics on this data, although I bear responsibility for the ultimate interpretation. The definitions follow from Howell (1999).

ratios calculated in section 4.2 will be the same in words with a weak-strong prominence pattern as the words with a strong-weak prominence pattern.

The *critical value* is the value set by the experimenter at which the $H_0$ will be rejected. In this study, the value that we will be looking at is the *p-value* (probability value). The p-value tells the probability that the results would occur by chance if the $H_0$ were true. In other words, the p-value tells us how likely it is that the difference (or relationship) found in the sample is not an actual difference that would be found in the population. The p-value is measured from zero to one, with zero meaning that it is impossible for the sample difference to be a reflection of the population difference, and one meaning that the sample difference is absolutely a reflection of the population difference. Thus, a p-value of .05 says that there is a five percent chance that the results of the sample data are not a reflection of the population data, and the difference occured by chance, or was a result of a sampling error. The critical value for this experiment will be expressed in terms of a p-value and will be set at .05. In other words, for this experiment to find a difference statistically significant, there must be a five percent chance (or less) that the difference did not occur by chance.

There will be two statistical procedures used in this study, ANOVAs and t-tests. *ANOVA*s (analysis of variance) look at the differences between groups and determine if this difference is significant by comparing it to the difference within groups. This study will use an ANOVA for two independent samples to determine whether the variance between the eight groups of different pitch and length patterns is larger than the variance within the groups. A separate ANOVA will be done for each of the five dependent variables. An ANOVA will give us a p-value to tell us the significance of the differences in the values between any of the eight groups. If the p-value is .05 or less, there will be a significant difference between at least two of the groups, but the ANOVA will not tell us which groups result in a significant difference.

To determine which groups have a significant difference between them, we turn to *t-tests*. T-tests give us the p-value for whether or not there is a significant difference between two groups. A separate t-test must be done each time you want to compare two groups. If there are five groups in an ANOVA and you want to see if there is a significant difference between each of the groups, 5! (or 120) t-tests will need to be done. Usually only a few of the t-tests will be done by the experimenter, depending on which groups are to be compared. T-tests can be one-tailed or two-tailed. A one-tailed t-test rejects the $H_0$ only if the sample is different from the population in one direction, (either specifically higher or specifically lower). In a two-tailed t-test it does not matter whether or not the sample is higher or lower than the population.

ANOVAs and t-tests will be done for independent samples. That is, the different groups of independent variables are not related to each other. Both

ANOVA and t-tests take into consideration thc number of items in the sample data. A higher number of items in the sample data will increase the chance that a significant difference will be found if there is one.

### 5.2. Results

A one-way ANOVA for independent samples was performed on the $F_0$ peak ratios on SPSS. The ANOVA looked at the variance between the eight groups based on length and prominence patterns. The p-value obtained was .0024. This tells us that there is a 0.24% likelihood that the different $F_0$ peak ratio values between the eight groups is a result of chance. This is below the critical value set for this experiment.

Because the different length patterns do not affect the $F_0$, the data will be split into two groups instead of eight for the t-tests: those with a strong-weak prominence pattern, and those with a weak-strong prominence pattern. A two-tailed t-test for independent samples was done by hand.[13] The result was a p-value less than .001. This tells us that there is less than a 0.1% chance that the different $F_0$ peak ratio values between strong-weak and weak-strong word occurred by chance. This is below our critical value of .05, thus we can say that $F_0$ is a significant correlate of prominence in Blackfoot. We can therefore conclude that Blackfoot speakers manipulate pitch deliberately in the marking of a prominent syllable.

A one-way ANOVA for independent groups was run for duration between the eight groups of differing length and prominence patterns. The p-value obtained was .0016. There is less than 0.16% likelihood that the differences in duration are a result of a sampling error. This is below our critical value of .05.

Unlike $F_0$, length patterns are relevant for the duration ratio measurements. T-tests will have to be performed between the two prominence patterns within the same length patterns. Short-short and short-long words were tested, as the other two length patterns have such small sample sizes. A two-tailed t-test for independent samples on words with a short-short pattern gives a p-value greater than .50. This says that there is over 50% likelihood that differences in total amplitude ratios between strong-weak and weak-strong words are due to chance. This is much higher than our critical value of .05. The $H_0$ that there is no difference in duration between the two prominence patterns in short-short words cannot be rejected, and thus duration cannot be said to be a correlate of Blackfoot prominence in these words.

A two-tailed t-test for independent samples on words with a short-long length pattern resulted in a p-value greater than .50. Again, this is much higher than our critical value of .05. The results of the t-tests suggest that there are no

[13] T-tests were not done on SPSS, they were calculated by hand. Exact p-values could not be determined, although ranges are given.

differences in the duration ratios between groups with different prominence patterns. The difference found in the ANOVA is likely due to the fact that Blackfoot has phonemic length. It would seem obvious that there would be difference between groups with different length patterns. Again, the $H_0$ cannot be rejected and according to these results, duration cannot be said to be a correlate of Blackfoot prominence in short-long words.

A one-way ANOVA for independent samples was run on the amplitude peak ratios of the eight groups, and a p-value of .1390 was obtained. This says that there is a 13.90% likelihood that the difference in the sample ratios is not a reflection of an actual difference in the population ratios. This is above our critical value of .05, and we cannot reject the null hypothesis that states that there is no difference in peak amplitude between the eight groups. That is to say that we cannot conclude a difference in peak amplitude between any of the eight groups. No further tests are needed because the ANOVA tells us that we would not find a significant value between any of the groups. We cannot conclude that peak amplitude is a correlate of prominence in Blackfoot.

The results of a one-way ANOVA for independent variables on the average amplitude ratios between the eight groups give a p-value of .0001. This is below our critical value of .05, and suggests that there is a difference between some of the groups. To determine which groups show different average amplitude ratio values, we must turn to t-tests.

A two-tailed t-test for independent samples was performed (like $F_0$, the length patterns are not relevant to average amplitude, therefore the data was split only into two groups, those with weak-strong and strong-weak prominence patterns). The p-value found was less than .001. This is well below the critical value of .05. This tells us that there is a less than 0.1% likelihood that the difference in average amplitude between the words with different prominence patterns is due to chance. Average amplitude is a significant correlate of prominence in Blackfoot. A prominent syllable in Blackfoot is deliberately marked with a higher average amplitude value; that is, it is louder overall than a syllable that is not prominent.

A one-way ANOVA for independent samples was done on the total amplitude ratios between the eight groups with different length and prominence patterns. The p-value obtained was .0027, which tells us that there is less than a 0.27% likelihood that variability found in the ratios between the eight groups is due to chance. This is below the critical p-value of .05, so t-tests will be done to see if the significant values are between groups with different prominence patterns.

Like duration, length patterns are relevant for total amplitude since duration is a factor in its measurement. T-tests were performed between the two prominence patterns within the same length patterns (short-short and short-long).

A two-tailed t-test for independent samples on words with a short-short pattern gives a p-value between .20 and .30. This says that there is between 20% and 30% likelihood that differences in total amplitude ratios between strong-weak and weak-strong words are due to chance. This is much higher than our critical value of .05.

A two-tailed t-test for independent samples on words with a short-long length pattern resulted in a p-value between .40 and .50. Again, this is much higher than our critical value of .05.

We know that a smaller sample size makes it less likely that a significant p-value will be found. It is a reasonable question to ask whether it is the difference in sample size in the total amplitude and duration t-tests that resulted in higher p-values as compared to those found in the $F_0$ and average amplitude ratios (because the total amplitude sample had to be split into groups with smaller sample sizes based on length patterns). To determine this, t-tests were run for the $F_0$ and average amplitude ratios between the prominence patterns of short-short and short-long words. Two-tailed t-tests on words with a short-long length pattern for the $F_0$ peak ratio and average amplitude ratio both produced p-values between .10 and .20. This is much lower than the duration value above .50, and also lower than the total amplitude value between .30 and .40. Although it is still not significant, we did find a significant value between strong-weak and weak-strong words when length patterns were not taken into consideration.

Two-tailed t-tests on words with a short-long length pattern for the $F_0$ peak ratio and average amplitude ratio were less than .01 and .001 respectively. This is much lower than the value obtained for duration that was above .50, and the value obtained for total amplitude that fell between .20 and .30.

These results show us that the difference in sample size in both the duration and total amplitude t-tests that resulted in higher p-values as compared to those found in the $F_0$ and average amplitude ratios is not a result of sample size. The null hypothesis that there is no difference in duration and total amplitude values in words with a strong-weak compared to a weak-strong prominence pattern cannot be rejected. Neither duration nor total amplitude can be said to be a correlate of Blackfoot prominence.

The results of the statistical analyses show that $F_0$, and average amplitude are correlates of Blackfoot prominence. That is, pitch and loudness over the whole syllable are deliberately manipulated as markers of word prominence by Blackfoot speakers. Table 12 is a summary table of the statistical results. The values that are not significant are shaded in.

**Table 12: Results of statistical analyses**[14]

| | ANOVA | T-tests | | |
|---|---|---|---|---|
| | | All groups | Short-Short | Short-Long |
| $F_0$ | p = .0024 | p < .001 | [redacted] | p < .01 |
| Duration | p = .0016 | n/a | [redacted] | [redacted] |
| Amp Peak | [redacted] | n/a | n/a | n/a |
| Ave Amp | p < .0001 | p < .001 | [redacted] | p < .001 |
| Tot Amp | p = .0027 | n/a | [redacted] | [redacted] |

## 6. General Discussion and Conclusions

### 6.1. Results of Acoustic and Statistical Analyses

The results of the analyses in this study suggest that $F_0$ peak and average amplitude are both correlated with prominence in Modern Blackfoot, while duration, amplitude peak and total amplitude are not. The correlation of the $F_0$ peak supports the assertion by Frantz (1989, 1991) and Kaneko that Blackfoot is a pitch accent language. This corresponds to Beckman's analysis of pitch accent in Japanese. As stated in §2.4, Beckman found $F_0$ peak was correlated with pitch accent in Japanese and all of the variables, primarily total amplitude, to be correlated with phonetic stress in English. The absence of a correlation to duration, average amplitude and total amplitude is similar to the Japanese results found by Beckman, but the correlation of amplitude peak differs. According to Beckman, amplitude peak is a correlation of phonetic stress and not pitch accent.

As discussed in §2.4, Sluijter and van Heuven say that amplitude is not a correlate of phonetic stress, suggesting that in Beckman's study, increases in amplitude were a result of the intonational contours overlaid on the phonetic stress. Sluijter and van Heuven also suggest that increases in amplitude associated with phonetic stress are a result of higher intensities at higher frequencies, while intensities at lower frequencies are affected very little. They found this to be a strong correlate of phonetic stress in Dutch, which was second only to duration.

It is difficult to determine what the reasons are for the correlations to prominence found in this study with average amplitude. Because this was one of the correlates Beckman found to phonetic stress, it could be suggested that Blackfoot has some characteristics of phonetic stress. Before coming to this conclusion, we must remember that Blackfoot was not found to have Beckman's most important cue of phonetic stress, total amplitude, or Sluijter and van Heuven's most important cue – duration. We must also remember that pitch,

[14] Shading indicates values that are not significant.

which was found to be a correlation of prominence in Blackfoot might not be a correlate of phonetic stress, but simply a result of intonation according to Sluijter and van Heuven. The results on Chuvash prominence found by Dobrovolsky (1999) also support this assertion. Because the words in this study were not elicited in a carrier phrase, it is difficult to determine what role intonation played on the dependent variables.

It may also be that Blackfoot is not like Japanese, as a prototypical pitch accent language, or English and Dutch, as prototypical phonetic stress accent languages, but falls somewhere in between the two methods of marking prominence. The patterns in §4.2.1 and §4.2.4 suggest that this is the case. Both variables show higher values on the first vowel in words with a strong-weak prominence pattern and similar values for the two words with a low high prominence pattern. This suggests that they are both correlates of word prominence, but also that they are both interacting with intonation. It would be interesting to see if the changes in amplitude peak in Blackfoot are a result of changes in Sluijter and van Heuven's spectral balance.

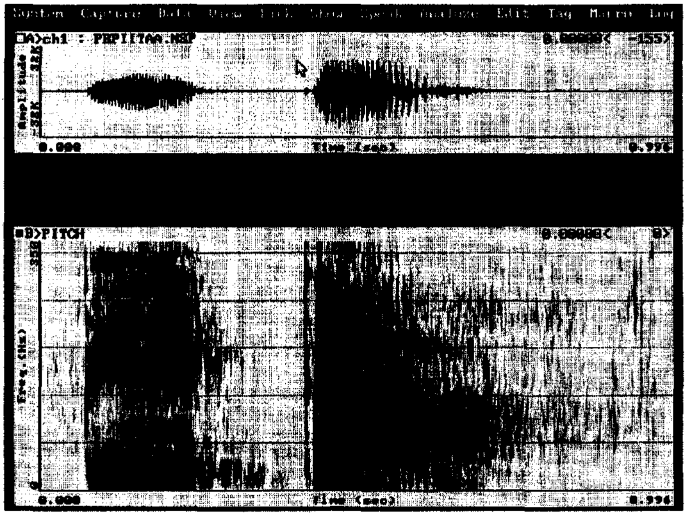
As discussed in §3.3, Kaneko claims that there are three different prominence patterns found in long vowels, those with a rising pitch, those with a falling pitch, and those with a high level pitch. Examples of these three different patterns were given in (14), restated here as (16).

| | | | |
|---|---|---|---|
| (16) | a. [pií.ta] | piíta | 'eagle' |
| | b. [póo.ka] | póoka | 'child' |
| | c. [aa.kíí] | aakíí | 'woman' (Kaneko 1999: 142-143) |

All three of the above words were examined in this study. As we have already seen, pitch is a correlate of prominence in Blackfoot, so we will look at the pitch patterns in these three words to test Kaneko's claim. Figure 16 shows the word *piítaa*[15] as given in (16a).

[15] The different length patterns in the vowels may be due to dialectal differences.

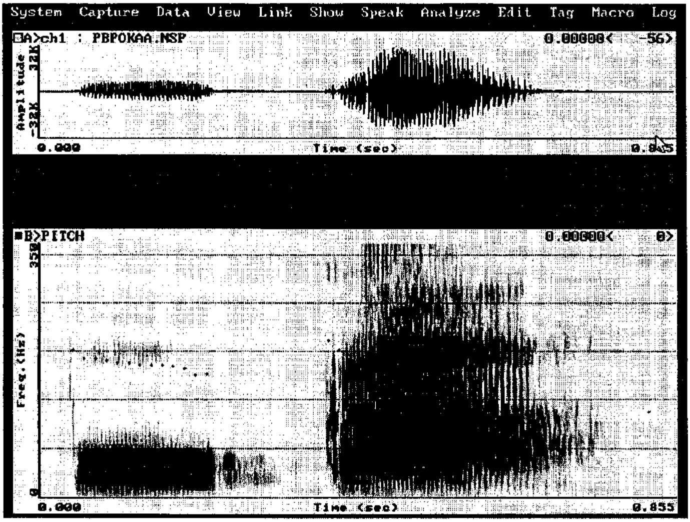
**Figure 16:** ***piítaa*****, 'eagle' (PB)**



You can see a definite rising pitch in the first syllable of *piítaa*, which is what was expected according to the form seen in (16a). This rising pitch in the first syllable goes against the falling pitch found in the syllables of all the other words. This was the only word examined in this study that displayed this rising pattern.

Figure 17 shows the pitch patterns for the word *póokaa*.

**Figure 17:** ***póokaa*****, 'child' (PB)**

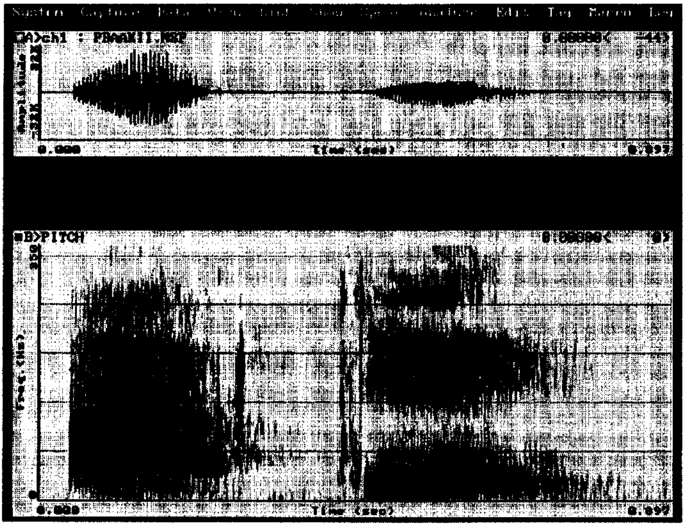



Although there is a slight falling pattern in the first vowel of *póokaa,* this is similar to the falling pattern found in almost all syllables (see Figure 2 in §4.1). It is not obvious that there is a falling pattern of prominence in *póokaa.* It is difficult to determine if there is an actual falling word prominence on any of the long vowels examined in this study as all syllables (except for the first syllable of *piítaa*) showed a falling pattern.

The pitch patterns for (16c), *aakíí,* can be seen in Figure 18. According to Kaneko, *aakíí* should have a level pitch on the final syllable. Although there is a falling pitch pattern, the actual word prominence is likely level throughout the long vowel. The falling pitch is due to the phrase final intonation. This pattern of falling pitch was found in all words looked at, regardless of vowel length.

All of the words with prominent long vowels (there were a total of fifteen) had a level or slightly falling pitch on the prominent syllable except for *piítaa.* Unfortunately, one word is not enough to give us evidence of different pitch patterns in long vowels. It may be that a level pitch pattern is much more frequent, while the rising and falling pitch patterns are not common. This could explain why only one word out of fifteen showed a definite contour pitch. Further analysis of pitch patterns in Blackfoot long prominent vowels is needed before this can be determined.

**Figure 18: *aakíí*, 'woman' (PB)**



### 6.2. Summary

The results from this study indicate that Frantz' (1989, 1991) and Kaneko's assertions that Blackfoot is a pitch accent language are correct. The acoustic evidence in this paper support this view. Pitch is a strong correlate of prominence in Blackfoot, just as it is in Japanese. The appearance of average amplitude as a correlate may be a result of intonation, or it may be that it is used as a cue for word prominence in Blackfoot. This latter possibility would suggest that Blackfoot pitch accent has acoustic differences from Japanese pitch accent, but I do not believe that it should constitute a different classification of prominence marking for Blackfoot.

### 6.3. Suggestions for Further Research

Further analyses would be useful on how amplitude patterns with word prominence and intonation in Blackfoot. Eliciting minimal pairs in carrier phrases as Sluijter and van Heuven did with Dutch would be useful in separating the intonation patterns from the pitch accent. An analysis of amplitude at higher frequencies would be useful in determining whether or not spectral balance is a correlate of prominence in Blackfoot. Elicitations of more words with a prominent syllable on a long vowel would give more evidence as to whether or not Blackfoot long vowels have different pitch patterns.

**Acknowledgements**
I would like to thank my supervisor, Michael Dobrovolsky for his patience, guidance and suggestions. I would also like to thank my second reader, John Archibald for his helpful comments. Thanks also to Suzanne Urbanczyk for introducing me to the Blackfoot language. And a special thanks to Pat and Noreen Breaker for sharing their language.

**Appendix A: Blackfoot Syntax and Morphology**
Blackfoot is a polysynthetic language with complex verbal and noun morphology. Verbs are inflected for subject, object and tense. Subject and object inflections show up as agreement morphology on the verb, and as a result, the pronouns they agree with may be omitted.

Nouns are inflected for number. Old Blackfoot[16] distinguished between animate and inanimate nouns, but this distinction seems to be being lost in Modern Blackfoot. For example, Frantz (1991:8) notes,

> A singular animate gender noun has -wa[5], and plural animate gender has -iksi
>
> [5] Certain speakers omit the suffix -wa under as yet undetermined conditions. And many young speakers never seem to use it.

This quotation shows that omitting *-wa* as a gender marker was already in process during Frantz' work. This is likely a result of a phonological phenomenon that is occurring in Blackfoot. Often the last syllable of a Blackfoot lexeme is devoiced, and sometimes deleted completely. This process will be discussed further in §3.1.2. Sometimes the same lexical item would be spoken both with a devoiced final syllable and again with the same syllable deleted completely. In fact, speaker B sometimes used the suffix *-wa*, although she indicated that it made no difference to the meaning when it was used. This may indicate that there are still some traces of gender distinction within Modern Blackfoot. In fact, when eliciting the form for "What is it?" speaker B said that there was a difference dependent upon whether or not the object was animate or inanimate. In this case, the gender is within an object agreement marker on the verb and is not the final syllable of the lexeme. This demonstrates how a phonological phenomenon is affecting the morphology and ultimately word classes within the language.

Modern Blackfoot has a basic SVO word order; however, there are some exceptions. The transitivity of the main verb within a Blackfoot sentence has a

[16] For simplicity, in this section only, Old Blackfoot will refer to the variety described by Frantz (1991) and Frantz & Russell (1989), and Modern Blackfoot will refer to the variety spoken by the consultants. This is the distinction used by the consultants when asked about the differences between the forms.

large affect on the morphology and syntax of a Blackfoot phrase. In some instances the word or morpheme order is changed as a result.

## Appendix B: Blackfoot Phonology

There is a relationship between the oral stops /t k/ and the corresponding affricates. The phoneme /t/ never occurs before the high-front vowel /i/, while /ts/ is very common in this environment. The affricate [ts] seems to be an allophone of /t/; /t/ becomes [ts] before /i/. However, this affricate also occurs in all of the environments that /t/ occurs in, which is why it is included in the phonemic inventory. The velar stop /k/ can occur before /i/, but it may change from [k] to [ks] in the same environment. This can be seen in (17).

| | | | |
|---|---|---|---|
| (17) | /sik-/ + /-i.ka/ | sik- + -ika | 'black' + 'foot' |
| | /si.ksi.ˈkaa/ | Siksika | 'Blackfoot' (Kaneko 1999: 156) |

This affrication may be due to an analogy with the relationship between [t] and [ts]. Because /t/ is realized as [ts] before /i/, /k/ may be realized as [ks] in the same environment.

It also appears that there is a phonological relationship between [x] and [k]. [x] only occurs in a coda position, and quite often it occurs before a [k]. Kaneko suggests that [x] is an allophone of /k/, and does not include /x/ in her phonemic inventory of Blackfoot. She gives convincing arguments for her proposal. She states that a single [k] never occurs in a coda position. This suggests that [x] and [k] are in complementary distribution and are allophones of a single phoneme /k/. However, one has to wonder what Kaneko means when she says that a *single* [k] never occurs in a coda position. Presumably, this is opposed to a geminate [k]. However, in geminates, the second half of the geminate would be syllabified as the onset of the following syllable while the first half remains in the coda position of the preceding syllable. In this sense, there is a single [k] in the coda. This can be seen in the data in (18) from Kaneko.

| | | | |
|---|---|---|---|
| (18) | a. [ni.ˈtak.kaa] | nitákkaa | 'my friend' |
| | b. [ˈpak.kiip] | pákkiip | 'choke cherry' (Kaneko, 1999:20) |

A similar pattern can be seen in (19), which shows data elicited from the consultants in this study. The gemination of the stop portion of the affricate /ks/ results in a /k/ in the coda.

| | | | | |
|---|---|---|---|---|
| (19) | [ˈksɪk.ksɪ.naa.tsi] | /ˈksik.ksi.naa.tsi/ | ksíkksinaatsi | 'white' (PB) |

In addition, elicitations with PB produced words with [k] in a word final coda position in a cluster before /s/. This is demonstrated in (20).

(20) [po.ˈkʊ.nɪks] /po.ˈko.niks/ pokóniks 'balls' (PB)

The absence of [k] in the coda position in environments other than gemination may at first seem to support Kaneko's view that [x], which frequently occurs in the coda position, is an allophone of /k/. However, the distribution of [t] and [p] is similar to that of [k]. They also do not occur in the coda position, except in environments similar to those found in (18) to (20). This may be due to constraints on the coda position in Blackfoot. It may also be that these phonemes can occur in Blackfoot codas, but only in low frequency due to the rules of syllabification in Blackfoot. If a non-geminate stop is intervocalic, it will be syllabified as an onset of the following syllable. If a stop is followed by an /s/, it will also be syllabified as part of the onset of the following syllable, reanalyzing the /ks/, /ts/ or /ps/ sequence as an affricate, which are phonemes of Blackfoot.

Although /x/ and /k/ do not seem to be allophones of the same phoneme, there does still appear to be some relationship between them. /x/ frequently appears in the coda position of syllables preceding a /k/ in the onset. The words in (21) demonstrate this pattern.

(21) a. [ˈum.ʔʌ̥x.kiː.na] /ˈum.**ʔ**ax.kiː.na/ ómahkiina 'old man' (PB)
b. [ˈsʌx.ko.maa.pi] /ˈsaax.**ko**.maa.pi/ sááhkomaapi 'boy' (PB, NB)

I could not find evidence of /x/ occurring as an onset. Kaneko's analysis of /k/ becoming [x] in the syllable coda suggests that her data also showed no evidence of [x] in the syllable onset position. I propose that a /k/ is inserted after an intervocalic /x/ in order to satisfy a constraint against /x/ in an onset position. A more thorough analysis of the distribution of these phonemes is needed to determine if this is truly the case. Until then, the relationship between /k/ and /x/ remains an issue that warrants further investigation.

There is one consonant in Kaneko's phonemic inventory that I have not included in my analysis, the glottal fricative /h/. She states that its distribution is very restricted, only occurring at the beginning of a few interjections. This phoneme was never encountered during elicitations with our native speakers. However, this may simply be a result of its limited distribution.

## Appendix C: Word List

### Weak-Strong Long-Long

| | | | |
|---|---|---|---|
| /aa.ˈkii/ | aakíí | 'woman' | (PB & NB) |
| /ii.ˈnii/ | iiníí | 'buffalo' | (PB) |
| /poo.ˈkaa/ | pookáá | 'child' | (PB & NB) |

### Weak-Strong Short-Long

| | | | |
|---|---|---|---|
| /ma.ˈmii/ | mamíí | 'fish' | (PB & NB) |
| /ox.ˈkii/ | ohkíí | 'water' | (PB & NB) |
| /piʔ.ˈksii/ | pi'ksíí | 'bird' | (PB) |

### Weak-Strong Short-Short

| | | | |
|---|---|---|---|
| /ki.ˈtsim/ | kitsím | 'door' | (PB) |
| /ni.ˈtan/ | nitán | 'my daughter | (PB) |
| /noʔ.ˈtsis/ | no'tsís | 'my arm/hand' | (PB & NB) |
| /po.ˈkon/ | pokón | 'ball' | (PB & NB) |

### Weak-Strong Long-Short

| | | | |
|---|---|---|---|
| /miis.ˈtak/ | miisták | 'mountain' | (NB) |
| /aa.ˈksin/ | aaksín | 'bed' | (PB) |

### Strong-Weak Long-Long

| | | | |
|---|---|---|---|
| /ˈpii.taa/ | piítaa | 'eagle' | (PB & NB) |

### Strong-Weak Short-Long

| | | | |
|---|---|---|---|
| /ˈni.naa/ | nínaa | 'man/chief' | (PB & NB) |
| /ˈpo.kii/ | pókii | 'child' | (PB & NB) |
| /ˈspi.taa/ | spítaa | 'tall' | (PB) |

**Strong-Weak Long-Short**

| | | | |
|---|---|---|---|
| /ˈkoo.si/ | kóósi | 'cup' | (PB) |
| /ˈnaa.ma/ | nááma | 'gun' | (NB) |
| /ˈnoo.ma/ | nóóma | 'my husband' | (PB) |

**Strong-Weak Short-Short**

| | | | |
|---|---|---|---|
| /ˈes.spi/ | áísspi | 'to dance (durative)' | (PB) |
| /ˈejnx.ki/ | áíynhki | 'to sing (durative)' | (PB) |
| /ˈaps.si/ | ápssi | 'an arrow' | (NB) |
| /ˈis.ska/ | ísska | 'pail' | (PB) |
| /ˈko.kit/ | kókit | 'to give' | (PB) |
| /ˈnin.na/ | nínna | 'my father' | (PB & NB) |
| /ˈnin.sta/ | nínsta | 'my sister' | (PB) |
| /ˈox.ki/ | óhki | 'it's barking' | (NB) |
| /ˈok.je/ | ókyai | 'hello' | (PB) |

## References


Beckman, M. E. 1986. *Stress and Non-Stress accent.* Dordecht-Holland/Riverton-USA:Foris.

Crystal, David. 1985. *A Dictionary of Linguistics and Phonetics: 2nd Edition.* Oxford:Basil Blackwell Ltd.

Dobrovolsky, M. 1999. The phonetics of Chuvash stress: Implications for phonology. In Ohala, John J., Yoko Hasegawa, Manjari Ohala, Daniel Granville and Ashlee C Bailey (eds) *Proceedings of the XIVth International Congress of Phonetic Sciences.*539-542.

Frantz, Donald G. 1991. *Blackfoot Grammar.* Toronto:University of Toronto Press.

Frantz, Donald G. and Norma Jean Russell. 1989. *Blackfoot dictionary of stems, roots, and affixes.* Toronto:University of Toronto Press.

Fry, D.B. 1958. Experiments in the perception of stress. *Language and Speech.* 1:126-152.

Howell, David C. 1999. *Fundamental Statistics for the Behavioral Sciences: 4th Edition.* Pacific Grove: Duxbury Press.

Hymen, Larry M. 1975. *Phonology: theory and analysis.* USA:Holt, Rinehart and Winston.

Kaneko, Ikuyo. 1999. *A metrical analysis of Blackfoot nominal accent in optimality theory*. MA dissertation. University of British Columbia.
Ladefoged, Peter. 1975. *A Course in Phonetics (second edition)*. Harcourt Brace Jovanovich.
Maclean, John. 1898. The Blackfoot language. *Canadian Institute Transactions*. 5:128-165.
O'Grady, William, John Archibald, Mark Aronoff and Janie Rees-Miller (eds). 2001. *Contemporary Linguistics, An Introduction: Fourth Edition*. USA:Bedford/St.Martin's.
Pike, Kenneth L. 1948. *Tone Languages*. Ann Arbor: University of Michigan Press.
Proulx, Paul. 1989.A sketch of Blackfoot historical phonology. *International Journal of American Linguistics*. 55:43-82.
Sluijter, Agaath M.C. and Vincent J. van Heuven. 1996. Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*. 100:2471-2485.
Uhlenbeck, C.C. 1938. *A Concise Blackfoot Grammar Based on Material From the Southern Peigans*. Amsterdam:Koninklijke Akademie van Wetenschappen, Verhandelingen, Afdeeling Letterkunde. n.s. 41.

Sheena Van Der Mark
Department of Linguistics
University of Calgary
2500 University Drive NW
Calgary, AB
T2N 1N4
Canada
dmscvan@ucalgary.ca

# *CWPL* STYLE SHEET

Documents should be submitted as camera ready hardcopies in accordance with the requirements outlined below. A copy should also be submitted on disk. **The editors reserve the right to return any submissions which do not adhere to the style sheet herein.** Note that this Style Sheet is also formatted according to these guidelines.

## 1 Manuscripts on disk

Manuscripts may be produced on a Macintosh computer in the following format using Microsoft Word or on a PC using Microsoft Word or WordPerfect. For tree structures, both ArborWin and MoraicWin are acceptable. Disk format required is **3.5 inch** and **high density**. If this is not possible, please contact the editors regarding alternate arrangements (including, but not limited to, e-mail submissions). If the disk is to be returned, a self addressed envelope is also required.

## 2 Hardcopy Manuscripts

**2.1** Hardcopy formatting (i.e., on paper) must be identical to the formatting on the disk.

**2.2** Manuscripts of articles submitted should be printed using laser quality print to ensure the maximal quality for copying. These copies will not be returned. Authors should retain hardcopies of their manuscript in their files.

**2.3** Manuscripts should be printed on letter size (8 1/2 x 11") paper on one side of the page only. All material, including extended quotes, footnotes, references, etc., should be single spaced, with double spacing at major divisions.

**2.4** Papers should not include page numbering. Authors are, however, asked to lightly write the page numbers on the **top right** corner of the **back** of each manuscript page in **pencil.**

**2.5** Left, right, top and bottom **margins** should be **3.5cm** (1.5").

**2.6** All text should be composed using **Times, IPA Times or IPA Extended Times** font. The size of the font should be 12 point for the text, 10 point for the footnotes and 7 point for the footnote numbers.

## 3 Manuscript Conventions

**3.1** All material, including extended quotes, footnotes, references, etc. should be single spaced except for indented quotes and examples which should have a single space between each examplar.

**3.2** Each article should begin with a blank line, the title on the second line, name of the author on the third line, and institutional affiliation or place of residence on the fourth line. This is followed by two blank lines then the abstract as per subsection 4.3, below. Titles should be short, descriptive, and straightforward and should be typed with Initial Caps. The author's name and affiliation should also be typed in Initial Caps.

**3.3** Footnotes, references, tables, diagrams, maps, etc. should be placed in their appropriate locations within the manuscript and not on separate pages.

**3.4** Section headings are required. Main headings should be **bolded** and underlined but not all-caps (e.g. **Introduction**). Section sub-headings are optional, but no more than one level of sub-headings should be used. Sub-headings should not be all-caps but should be **bolded** (e.g. **Sentence Types**). There should be **no space** between section headings and their accompanying text but there should be a **single space** between distinct sub-headings and a **double space** before a Main Heading.

**3.5** All text should be **fully justified** including abstracts, text body, footnotes, references, etc.

## 4 Text Conventions

**4.1** Linguistic forms cited within a sentence in the text should be set apart from the text. Recommended conventions are as follows:

(1) Forms cited in phonetic transcription should be enclosed between [square brackets].

(2) Forms cited in phonemic transcription should be between /slant lines/.

(3) Other cited forms (e.g., underlying forms) should be underlined.

(4) Authors may designate other transcription conventions such as {curly brackets} or ⟨obliques⟩ but must define their use.

**4.2** Glosses of single words or morphemes should be enclosed between single quotation marks, which are not otherwise used (e.g., /amihkw/ 'beaver'). Double quotation marks should be used only for short quotations, reported conversation and the like.

**4.3** Glosses of phrases or sentences should be placed below the example and aligned with the corresponding morpheme or word. An English translation should follow on the third line. This should be single-spaced and numbered serially with Arabic numerals as per subsection 4.6, below.

**4.4** Authors are asked to include an abstract of their paper under the title, their name and their institution. The title **Abstract** should be centred and bolded above the abstract. **A separate copy of the abstract should also be submitted with the paper to be sent to a publisher of Working Paper Abstracts.**

**4.5** The **abstract** and **extended quotations** of three typed lines or more should be set apart from the main text by **double spacing** both **before** and **after** the quotation, should be single spaced, and with both the left and right margins indented five spaces. No quotation marks of any sort should be used.

**4.6** Sets of examples or example sentences should be numbered serially with Arabic numerals closed in parentheses. If several such examples are grouped together, the entire group is identified by an Arabic numeral, and the individual sentences by lower case letters:

(3) a. John loves Mary.
b. Mary is loved by John.

Rules set off from the text should be similarly numbered:

(8) C --> [-voice]/______#

## 5 Table/Figure Conventions

**5.1** Number figures and tables consecutively (figures separately from tables) with Arabic numerals (i.e., **Table 3** or **Figure 6**). All figures and tables should be placed in their respective places within the text.

**5.2** A brief title for each table/figure that makes the data intelligible without reference to the text may be used. Longer explanatory material should be typed as a footnote to the table, not as part of the title.

**5.3** Column heads should be short, so as to stand clearly above the columns.

## 6 Footnote Conventions

**6.1** Footnotes should be located at the bottom of the page. They should be typed beginning with a raised number with a single space between each note.

**6.2** Footnotes are not used for bibliographical reference. They should be brief, ancillary comments on the main text and not extended discussions.

**6.3** Footnotes should be numbered consecutively throughout the text. A footnote number in the main text is to be typed as a raised number immediately following the material to which it refers, e.g.:

...the extended linkage[3] which is...

Footnotes at the end of a sentence should follow the final punctuation:

...as evidenced in Gothic.[3]

**6.4** Acknowledgements should be placed immediately after the text but immediately before the references. This should be titled **Acknowledgements** and should be left-aligned. Leave a double space after the text before the Acknowledgements and similarly between the end of the Acknowledgements and the References.

## 7 Reference Conventions

**7.1** Complete bibliographical information is not cited in the text or as a footnote. Within the text, the author's name, the date of the work referred to, and the page number(s) (if appropriate) are sufficient. The reference should be between parentheses:

... it has been suggested (Johnson 1959:32) that ...

If the author's name is part of the sentence, only the numbers are between parentheses:

... Johnson (1959:32; 1962:297) has suggested that ...

If the author's name is part of a parenthetical comment, the parentheses are omitted from the numbers, e.g.:

... some authors have suggested (Johnson 1959:32, Key 1960:98 and Smith 1963:65) that ...

**7.2** Do not use the terms "ibid." and "op.cit." Where necessary, to avoid ambiguity, repeat the full reference. Do not use authors' initials when citing references in the text unless necessary to distinguish two authors of the same surname.

**7.3** Full bibliographical information for the references cited in the text should be located within the section entitled **References** (left-justified) at the end of the paper. Entries should be single-spaced both within and between references. Works are listed alphabetically by author's last name, and chronologically when two or more works by the same author are listed, distinguished by lower case letters in the case of works published in the same year. Each entry has four elements: the author's name, the year published, the title, and the source or place of publication. Each line following the first line of an entry is indented eight spaces. Titles of books should be in italics. Titles of both books and articles should follow the convention where only the first word of the title is capitalised. All other words, with the exception of proper nouns, should be in lower case. The following patterns should be used:

Single author:

Sapir, Edward. 1921. *Language*. New York: Harcourt, Brace.

Single Editor:

Fishman, Joshua A., ed. 1968. *Readings in the sociology of language*. The Hague: Mouton.

Multiple authors:

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Articles:

Jasanoff, Jay. 1978. 'Observations on the Germanic Verschärfung.' *Münchener Studien zur Sprachwissenschaft*. 37:77-90.

If in doubt, follow the guidelines offered in *Language* or *Linguistic Inquiry*, insofar as they do not conflict with this Style Sheet.

## 8 Name and Address

Authors should include their name, address, fax number, and email address at the bottom of their paper following the **References**.

Example:

Bilbo Baggins
Department of Linguistics
University of Middle Earth
144 Bag End,
West Farthing, The Shire
Z1A 5L2 Eriador
bbaggins@ling.ume.er